CONNECTING GENES TO MOLECULES: IDENTIFYING SMALL MOLECULES FROM PHYTOPATHOGENIC BACTERIA

Erinn M. O'Neill

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill
in partial fulfilment of the requirements for the degree of Doctor of Philosophy
in the Department of Chemistry.

Chapel Hill
2019

Approved by:

Bo Li

Gary Pielak

Matthew Redinbo

Matthew Lockett

Leslie Hicks

ABSTRACT

Erinn M. O'Neill: Connecting genes to molecules: Identifying small molecules from phytopathogenic bacteria
(Under the direction of Bo Li)

Bacterial plant pathogens cause significant crop damage worldwide. They invade plant cells by producing a variety of virulence factors, including small molecule toxins and phytohormone mimics. Virulence of the model pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) is regulated in part by the sigma factor HrpL. Our study of the HrpL regulon identified an uncharacterized, three-gene operon in *Pto* that is controlled by HrpL and related to the *Erwinia hrp*-associated systemic virulence (*hsv*) operon. Here, we demonstrate that the *hsv* operon contributes to the virulence of *Pto* on *Arabidopsis thaliana* and suppresses bacteria-induced immune responses. We show that the *Pto hsv*-encoded enzymes synthesize a novel small molecule, phevamine A. This molecule consists of L-phenylalanine, L-valine and a modified spermidine, and is different from known small molecules produced by phytopathogens. We show that phevamine A suppresses a potentiation effect of spermidine and L-arginine on the reactive oxygen species burst generated upon recognition of bacterial flagellin. The *hsv* operon is found in the genomes of divergent bacterial genera, including ~40% of *P. syringae* genomes, suggesting that phevamine A is a widely-distributed virulence factor in phytopathogens. Our work identifies a new small molecule virulence factor and reveals a new mechanism by which bacterial pathogens overcome plant defense. The workflow from this project was also applied to two other phytopathogenic strains of *Pseudomonas*

syringae to identify novel natural products from *Pseudomonas syringae* pv. *syringae* UMAF0158 and *Pseudomonas syringae* pv. *actinidiae* NZV-13. Another small molecule, pepticinnamin E, was identified from *Actinobacteria bacterium* OK006 for the first time and efforts are underway to understand its biosynthesis. This work highlights the power of -omics approaches in identifying important small molecules in bacteria–host interactions.

To my parents for believing

in me and always pushing me

to try my hardest.

ACKNOWLEDGEMENTS

I would like to thank my parents, Patrick and Christine. You've pushed me to try my hardest in everything I've ever done. Any time I faced a challenge I didn't think I could overcome, you supported me in whatever choice I made and it means so much to me. I would also like to thank my sister, Caitlin, for teaching me everything she was learning in school even though I was five years younger. You've helped me become the quick learner I am today. I love you all. I would also like to thank my extended family, both alive and deceased, for shaping me into the person I am today and giving me lots of laughs along the way.

I would like to thank my advisor, Professor Bo Li. You believed in me when I was an analytical chemist without any biology background, and you gave me the tools to succeed. You challenged me with projects and helped me change how I think so I can finally see the bigger picture and not just small details anymore. I have become a well-rounded chemist all thanks to you. I absolutely would not be where I am today without you, thank you.

I would also like to thank my committee members for the support throughout the years, especially Dr. Gary Pielak for answering my many questions about graduate school. Thanks to Professor Jeffery Dangl, Dr. Tatiana Mucyn, Dr. Omri Finkel, and Dr. Eui-Hwan Chung for collaboration on plant biological assays for the phevamine project. Thanks to Professor Frank Schroeder and Dr. Joshua Baccille for collaboration on NMR analysis for the phevamine project.

I owe great thanks to my labmates, especially Andy Chan, Kevin Santa Maria, Zach Dunn, and Ashley Kretsch. We all came into Bo's lab at the same time and it has made us feel like family and has brought us very close together. Andy, you're one of the best friends I've ever had and you've helped me through so much – a very special thank you to you!! I love you. I would also like to thank Jon Patteson, Gina Morgan, Rachel A. Johnson, Katie Acken, Adam Lescallette, Rachel M. Johnson, Martina Knechel, and Savannah Weeks for lightening up the mood in lab and making work more fun. Rachel A. Johnson and Zach Dunn were also my roommates for a few years and I'm so grateful for all the fun times we had living together. I am also grateful for the former postdoctoral researchers in our lab, Dr. Jillian Tyrrell, Dr. Wenlong Cai, and Dr. Elisabetta Massolo for helping me whenever I had questions.

Lastly, I owe a huge thanks to my lifelong friends, who have supported me all throughout graduate school and I know will continue to support me the rest of my life. Kim Junkins, Steve Hasiak, Jarod Grossman, Nick Chiappini, Ramaya Farooq, Brigid Brosnan, Colleen Brosnan, Evan Amber, Stephan Icso, Ronak Mistry, Sarah Colmer, Michal Rudiak, John McHale, John Choffo, Brian Gallagher, you've all been great friends to me.

I am truly grateful for each of you, thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Adenylation
ACN	Acetonitrile
Ala	Alanine
AMP	Adenosine monophosphate
Arg	Arginine
Asn	Asparagine
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
C	Condensation
Cor	Coronatine
Cys	Cysteine
DAP	2,3-diaminopropionate
Dha	Dehydroalanine
Dhb	Dehydrobutyrine
DOPA	Dihydroxyphenylalanine
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
FA	Formic acid
Fmoc-Cl	9-Fluorenylmethoxycarbonyl chloride
FPLC	Fast protein liquid chromatography
Gln	Glutamine
Glu	Glutamate

Gly	Glycine
hArg	Homoarginine
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hpg	Hydroxyphenylglycine
HPLC	High-performance liquid chromatography
Hrp	Hypersensitive response and pathogenicity
Hsv	Hrp-associated systemic virulence
Ile	Isoleucine
KB	King's B media
LB	Lysogeny broth
LC	Liquid chromatography
LC-HRMS	Liquid chromatography-high resolution mass spectrometry
Leu	Leucine
Lys	Lysine
M9	M9 minimal media
MAMP	Microbe-associated molecular pattern
Met	Methionine
MS	Mass spectrometry
MS^2	Tandem mass spectrometry
MT	Methyltransferase
MTI	MAMP-triggered immunity
NMR	Nuclear magnetic resonance spectroscopy
NRPS	Nonribosomal peptide synthetases

OAT	Ornithine acetyltransferase
ODC	Ornithine decarboxylase
OCT	Ornithine carbamoyltransferase
ORF	Open reading frame
Orn	Ornithine
P_i	Free phosphate
PCR	Polymerase chain reaction
Pg	Phenylglycine
Phe	Phenylalanine
PHVA	Phevamine A
PHVB	Phevamine B
PKS	Polyketide synthase
PMS	*Pseudomonas* minimal salts media
pPHV	Prephevamine
PP_i	Pyrophosphate
Pro	Proline
Q-TOF	Quadrupole time of flight
R	Reductase
ROS	Reactive oxygen species
Ser	Serine
Spd	Spermidine
Spn	Spermine
T	Thiolation

T3SS	Type III secretion system
TFA	Trifluoroacetic acid
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
UV	Ultraviolet
Val	Valine
WT	Wildtype

CHAPTER 1: INTRODUCTION – ATP-GRASP ENZYMES IN NATURAL PRODUCT BIOSYNTHESIS

1.1 OVERVIEW OF ATP-GRASP ENZYMES

ATP-grasp enzymes share a highly conserved ATP-grasp fold that "grasps" a molecule of ATP. These enzymes are composed of an N-terminal N-domain, a central B-domain, a central A-domain, and a C-terminal C-domain (Fig. 1.1).[1] The ATP-grasp site is located between the C-terminus and the central domains, where ATP is bound and used for substrate activation.[1]

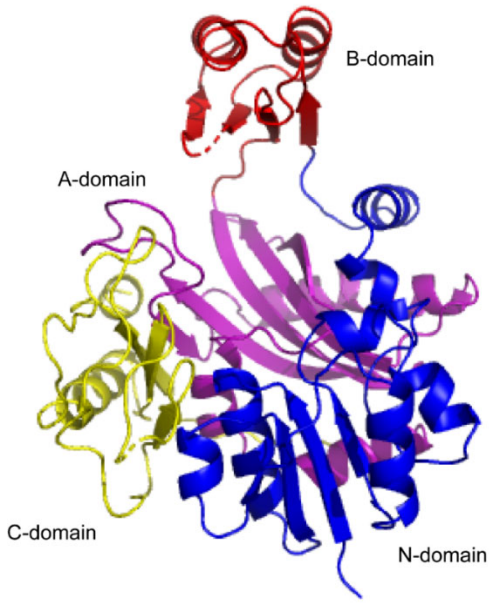


Figure 1.1. General structure of ATP-grasp enzymes. The X-ray crystal structure of *E. coli* glycinamide ribonucleotide synthetase (PDB 1GSO) is shown.[2] ATP-grasp enzymes contain three subdomains: an N-terminal N-domain (blue), a central B-domain (red), a central A-domain (purple) and a C-terminal C-domain (yellow). The ATP-grasp fold is located between the A- and C-domains.

The mechanism of ATP-grasp enzymes requires the carboxylic acid of one substrate to first be activated by one phosphate of ATP, thus forming a reactive acylphosphate intermediate and releasing ADP (Fig. 1.2).[1, 3] Once activated, the carbonyl carbon is electrophilic and can undergo nucleophilic attack by an amine substrate, forming an amide bond and then releasing free phosphate (P_i) (Fig. 1.2).[1, 3] This mechanism of amide bond formation is independent of the ribosome, and also differs from adenylation enzymes such as non-ribosomal peptide synthetases (NRPS).

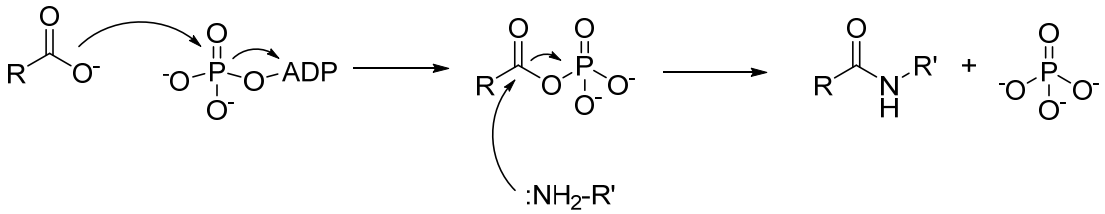


Figure 1.2. General mechanism of ATP-grasp-catalyzed amide bond formation. The carboxyl of one substrate is activated by phosphorylation, generating an acylphosphate intermediate. The amine of the second substrate then attacks the carbonyl of the acylphosphate, forming an amide bond and releasing free phosphate.

Many in-depth studies of ATP-grasp enzymes have focused on those involved in amide formation in metabolic pathways, including fatty acid biosynthesis, purine biosynthesis, and gluconeogenesis.[1, 4-8] The most well-studied ATP-grasp enzymes include D-Ala-D-Ala ligases in peptidoglycan biosynthesis, biotin carboxylases in fatty acid biosynthesis, and RimK in *E. coli* ribosome biogenesis.[4-5, 7-8] Many ATP-grasp enzymes are also involved in the biosynthesis of secondary metabolites or natural products. Bioinformatic studies revealed that microorganisms harbor many ATP-grasp enzymes in their genomes, most of which are uncharacterized.[3, 9] Most genome-mining efforts center on biosynthetic gene clusters

containing assembly line-like proteins, such as NRPS and polyketide synthases (PKS), and the clusters containing ATP-grasp enzymes have been mostly overlooked.

This chapter focuses on microbial ATP-grasp enzymes that synthesize natural products, including their chemical mechanisms, substrate scope, and structures, to understand the molecular basis for substrate specificity. It will also discuss recent genome-mining efforts for ATP-grasp enzymes to discover new natural products.

1.2 NATURAL PRODUCT PATHWAYS INVOLVING ATP-GRASP ENZYMES

ATP-grasp enzymes contribute to the biosynthesis of many natural products. These compounds exhibit diverse structures (Fig. 1.3), biological activities, and mechanisms of action, including phytotoxins, antibiotics, and antioxidants.

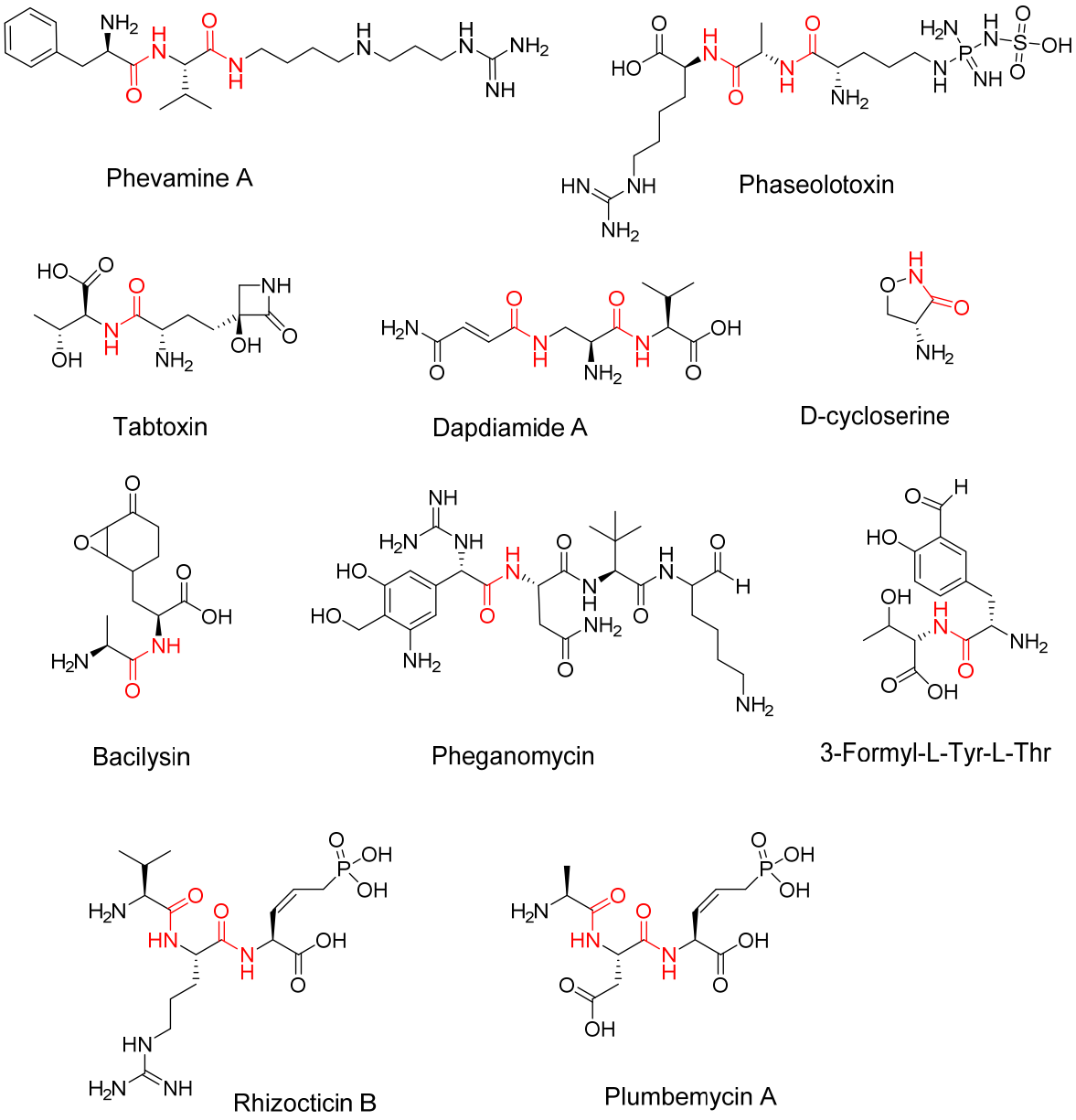


Figure 1.3. Chemical structures of several ATP-grasp synthesized natural products. ATP-grasp-catalyzed amide bonds are highlighted in red.

Tabtoxin is a phytotoxin from *Pseudomonas syringae* that inhibits glutamine synthetase. The ATP-grasp enzyme TabS catalyzes the condensation of L-threonine to the tabtoxin precursor, tabtoxinine-β-lactam.[10] TabS exhibits extremely broad substrate scope in vitro: it catalyzes formation of 136 other dipeptides using amino acids that have vastly different

chemical properties.[10] These dipeptides include: all Met dipeptides, Cys-Leu, Ser-Phe, Trp-Tyr, Glu-β-Ala, Pro-Thr, Asn-Gln, and Gly-Val.[10] TabS can incorporate every proteinogenic amino acid into at least one dipeptide, which suggests that the active site of TabS is highly plastic to accommodate both aromatic amino acids, like Trp and Tyr, and charged amino acids as substrates.

Another *P. syringae* phytotoxin is phaseolotoxin. This molecule contains a *N*-sulfodiaminophosphinyl-L-ornithine attached to L-alanine and L-homoarginine, and inhibits plant ornithine carbamoyltransferase (OCT) and ornithine decarboxylase (ODC).[11] The gene cluster encodes two ATP-grasp enzymes, PhtQ and PhtU, which have been proposed to catalyze Ala-hArg and Orn-Ala-hArg formation, respectively.[11] However, these enzymes have not been biochemically characterized and their substrate scope is unknown.

Phevamine A was identified by mining a *P. syringae* gene cluster encoding two ATP-grasp enzymes.[12] This cluster was upregulated by a sigma factor that controls a variety of virulence pathways in *P. syringae.*[12-13] In vitro reconstitution of the enzymes encoded in the gene cluster, including the ATP-grasp enzymes HsvB and HsvC, enabled the identification and structural characterization of phevamine A, which consists of L-phenylalanine, L-valine, and amidinospermidine.[12] Phevamine A was shown to promote bacterial virulence by suppressing plant immune response in *Arabidopsis thaliana* and *Nicotiana benthamiana.*[12] In phevamine A biosynthesis, HsvC ligates L-Val to amidinospermidine and HsvB ligates L-Phe to L-Val-amidinospermidine to form phevamine A.[12] Under in vitro and heterologous expression conditions, HsvB also links a second L-Phe to phevamine A, forming phevamine B.[12] In the absence of any cosubstrates, HsvB catalyzes formation of Phe-Phe dipeptide and Phe-Phe-Phe tripeptide.[12] Additionally, HsvB can only activate L-Phe but not D-Phe or L-Phe-L-Phe

dipeptide as substrate (unpublished data). These data suggest that HsvB selectively activates the carboxyl of L-Phe to form the acylphosphate intermediate, but accepts both single amino acids and peptides as the nucleophile. The substrate scope of HsvC remains to be explored.

The biosynthetic operon for mangotoxin shares similarity to that of phevamine A; it also consists of two ATP-grasp enzymes, MboC and MboD, and an amidinotransferase, MboE.[14] Mangotoxin was identified in 2003 from some strains of *P. syringae* as a phytotoxin.[15] It acts as an antimetabolite that inhibits plant ornithine acetyltransferase (OAT) and causes necrosis in plant leaves.[15] The structure of mangotoxin remains unknown and the functions of the biosynthetic enzymes including MboC and MboD are uncharacterized.

Dapdiamides are tripeptide antibiotics produced by *Pantoea agglomerans* that inhibit the growth of the phytopathogenic bacterium *Erwinia amylovora*.[16] Their biosynthesis requires two ATP-grasp enzymes, DdaF and DdaG.[16] DdaG catalyzes amide bond formation between fumarate and 2,3-diaminopropionate (DAP), and DdaF catalyzes the second amide bond formation between N_β-fumaramoyl-DAP and L-valine to form dapdiamide A, L-isoleucine to form dapdiamide B, or L-leucine to form dapdiamide C.[16] DdaF will not accept N_β-fumaroyl-DAP as a substrate, and the fumaroyl moiety must first be converted to the fumaramoyl half amide.[16] Further, DdaF will not accept N_α-fumaramoyl-DAP either, suggesting a narrow scope of DdaF for N-terminal substrate specificity.[16] Similarly, DdaG prefers the native fumarate substrate over fumaramate, succinate, (*R,R*)-epoxysuccinate, (*S,S*)-epoxysuccinate, and *cis*-epoxysuccinate on the N-terminus.[16] DdaF is more flexible toward the C-terminal substrate and incorporates the branched aliphatic amino acids Val, Ile, and Leu.[16] Further studies of DdaF with other amino acid substrates such as Ala or Thr would yield additional insights on the specificity of this enzyme.

Streptomyces lavendulae produces D-cycloserine, an aminoglycoside antibiotic that is active against tuberculosis.[17] This molecule has an interesting scaffold, which is in part due to the ATP-grasp enzyme DscG, which cyclizes o-ureido-D-serine.[17] This cyclization reaction is unique for an ATP-grasp enzyme, compared to the linear peptides this class of enzymes typically form. Interestingly, DcsG was able to catalyze formation of five- or six-membered ring products from the following linear substrates: D-homocysteine, β-aminooxy-D-alanine, D-homoserine, D-2,4-diaminobutyrate, and D-ornithine.[17] However, D-asparagine was not cyclized by DcsG.[17] The chemical properties of the sidechains in homocysteine, homoserine, and ornithine are very different, but all three are cyclized by DcsG, suggesting the substrate specificity of the enzyme may be influenced by size of the sidechain rather than electrostatic interactions between the sidechain and the active site of the enzyme.

The antioxidant shinorine from *Microcystis aeruginosa* is synthesized by MysC, which links glycine and 4-deoxygadusol, and MysD, which catalyzes the final step condensing the shinorine precursor to L-Ser (Fig. 1.4).[3, 18] Although both of these enzymes are ATP-grasp enzymes, neither forms an amide bond. Instead of serine, MysD has been shown to also act on glycine and threonine, but other amino acid substrates have not been tested.[18] Similarly, MysC substrate specificity has not been tested either. Since 4-deoxygadusol is a noncanonical substrate for ATP-grasp enzymes, MysC may have strict substrate control and perhaps a small active site since Gly is the other substrate.

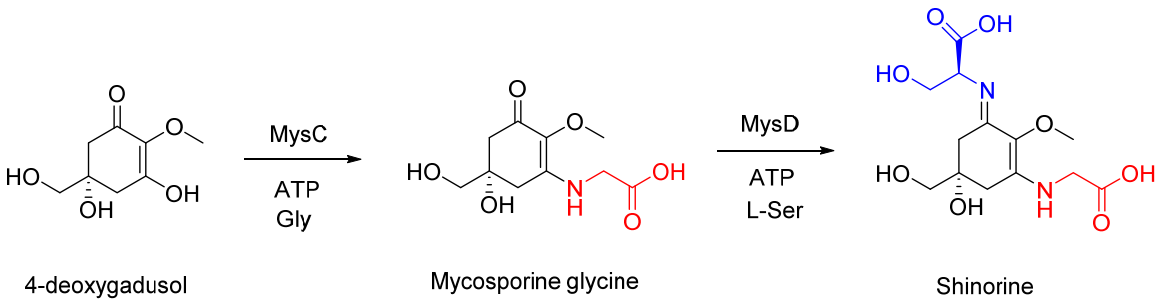


Figure 1.4. Biosynthesis of shinorine by the ATP-grasp enzymes MysC and MysD. The ATP-grasp enzymes MysC and MysD synthesize shinorine from 4-deoxygadusol, but neither form an amide bond. The addition of Gly to 4-deoxygadusol by MysC is shown in red, and the addition of L-Ser to mycosporine glycine by MysD is shown in blue.

Other natural products synthesized by ATP-grasp enzymes include: nikkomycins, friulimicin B, plumbemycins, argolaphos, and fosfazinomycin A.[3-4, 19] Nikkomycins are peptidyl nucleoside antibiotics produced by *Streptomyces tendae* and biosynthesis of nikkomycins requires the ATP-grasp enzyme NikS.[20] Insertional mutants in the nikkomycin pathway showed NikS is essential for production of nikkomycins I, J, X, and Z, however the substrate scope and structure of NikS remains to be explored.[20] Friulimicin B is a lipopeptide antibiotic produced by *Actinoplanes friuliensis* and is synthesized in part by the ATP-grasp enzyme DabC.[4, 21] However, DabC activity has not been reconstituted *in vitro.* Plumbemycins are tripeptide antibiotics produced by *Streptomyces plumbeus.*[22] The plumbemycin gene cluster has been elucidated and contains several ATP-grasp enzymes, but these enzymes have not been characterized.[23] Argolaphos A and B are synthesized in part by an ATP-grasp enzyme that has not been biosynthetically characterized.[19] These compounds are broad-spectrum antibacterial phosphonopeptides produced by *Streptomyces monomycini.*[19] Finally, the fosfazinomycin A pathway encodes the putative ATP-grasp enzyme FzmF, which has not been

biochemically characterized.[24] Fosfazinomycin A is produced by *Streptomyces* sp. WM6372 and is active against some filamentous fungi.[24-26]

ATP-grasp enzymes as a whole have robust substrate scope and contribute to the biosynthesis of many natural products. These enzymes are present in a diverse assortment of microbes and produce many varied natural product scaffolds.

1.3 STRUCTURES AND MECHANISMS OF ATP-GRASP ENZYMES IN NATURAL PRODUCT BIOSYNTHESIS

ATP-grasp enzymes are composed of four subdomains: N, A, B, and C. Between the ATP-grasp subdomains are several long loop regions, which allow for protein flexibility and recognition of structurally diverse substrates.[27] Crystal structures of ATP-grasp enzymes provide a structural basis for understanding ATP-grasp activity and substrate scope.

X-ray crystallography is a widely-used method for identification of crucial protein active site residues and characterization of macromolecular structure. Several ATP-grasp enzymes involved in secondary metabolite production have been crystallized, including MdnB and MdnC in the microviridin J pathway, PGM1 in the pheganomycin pathway, BacD in the bacilysin pathway, RizA in the rhizocticin pathway, and BL00235, which catalyzes dipeptide formation.

Microviridin J is a large macrocyclic ribosomally synthesized and post-translationally modified peptide (RiPP), composed of 13 modified amino acids.[27] The ATP-grasp enzyme MdnC catalyzes two lactonizations, forming the first two macrocycles of microviridn J, and MdnB catalyzes the final macrocycle formation (Fig. 1.5A).[27] Both MdnB and MdnC form homodimer assemblies with interface areas of 2,422 and 2,748 Å^2, respectively.[27] One conserved ATP-grasp α-helix, α3, extends along the central domain of the enzyme. In the

microviridin enzymes, this α-helix is 23 residues long, which is much longer than typical α3 helices.[27] However, most ATP-grasp enzymes involved in natural product biosynthesis act on substrates much smaller than RiPPs. This natural product presents an interesting case where MdnB and MdnC also possess a RiPP recognition element for recognition of the leader peptide on the peptide substrate.

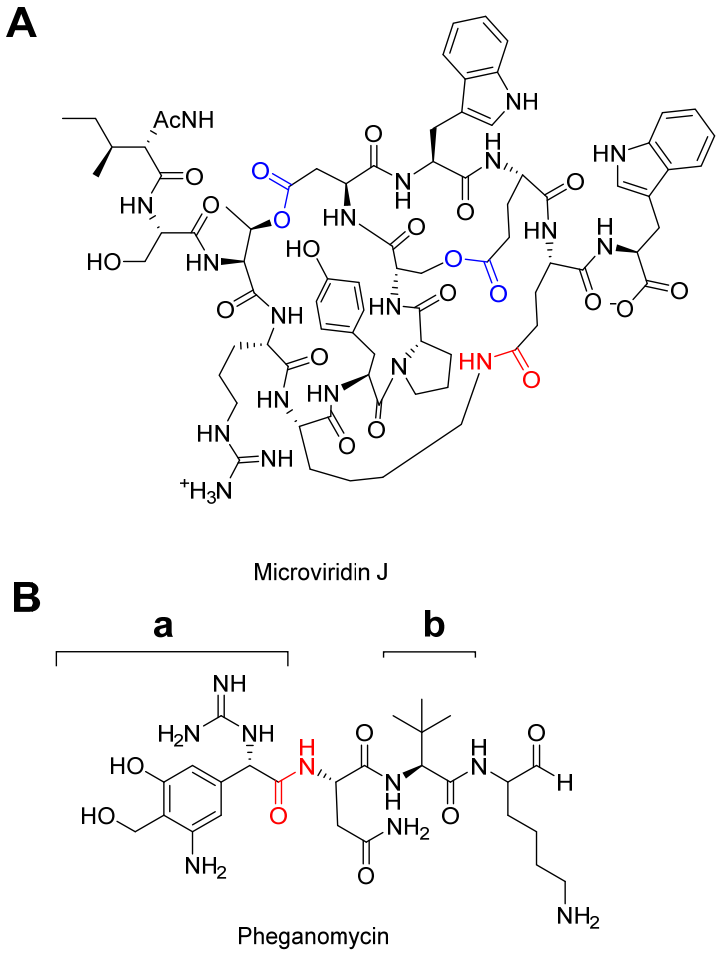


Figure 1.5. Chemical structures of large ATP-grasp synthesized secondary metabolites. (*A*) Structure of microviridin J. (*B*) Structure of pheganomycin showing the two nonproteinogenic amino acid components: **a**, (*S*)-2-(3,5-dihydroxy-4-hydroxymethyl)phenyl-2-guanidinoacetic acid, and **b**, (*S*)-2-amino-3,3-dimethylbutanoic acid (L-tert-leucine). ATP-grasp-catalyzed lactonizations are shown in blue and ATP-grasp-catalyzed amide bonds are shown in red.

PGM1 is responsible for peptide linkages to generate the linear peptide pheganomycin and related analogs (Fig. 1.5B).[28] Pheganomycin is an anti-*Mycobacterium* antibiotic produced by the yeast species *Saccharomyces cerevisiae.* The molecule is composed of two nonproteinogenic amino acids, (*S*)-2-(3,5-dihydroxy-4-hydroxymethyl)phenyl-2-guanidinoacetic acid (**a**) and (*S*)-2-amino-3,3-dimethylbutanoic acid (**b**), and two proteinogenic L-amino acids, asparagine and lysine.[28] PGM1 links the nonproteinogenic amino acid (**a**) to a core peptide.[28] The crystal structure obtained for PGM1 revealed a large cavity with a deep cleft in the center of PGM1 with the ATP-binding site located deep in this cavity (Fig. 1.6).[28] PGM1 therefore has a large active site to accommodate larger peptide substrates and may be very promiscuous, accepting many peptide substrates. PGM1 was found to to accept other nonproteinogenic amino acids with guanidino groups as the N-terminal substrate, including (*S*)-2-guanidino-2-phenylacetic acid, (*R*)-2-guanidino-2-phenylacetic acid, 2-guanidinoacetic acid, and creatine.[28] Conversely, (*S*)-2-amino-2-(3,5-dihydroxyphenyl)acetic acid, 3-guanidinopropanoic acid, 4-guanidinobutanoic acid, and all L-amino acids were unable to be linked at the N-terminus to the rest of the pheganomycin peptide by PGM1.[28] All L-amino acids were also unable to serve as C-terminal substrates.[28] The natural C-terminal substrates are the peptides NVKDR, NVKDGPT, NV, NVKDG, and NVKDGP.[28] Other accepted C-terminal substrates include AVKDR, NAKDR, NVADR, AAKDR, AAADR, aspartame, MRFA, GNNRPVYIPQPRPPHRPL, and RVRRPVYIPQPRPPHPRL.[28] This presents an interesting case, where single amino acids are not accepted substrates by the ATP-grasp enzyme, but large peptides and nonproteinogenic amino acids are accepted.

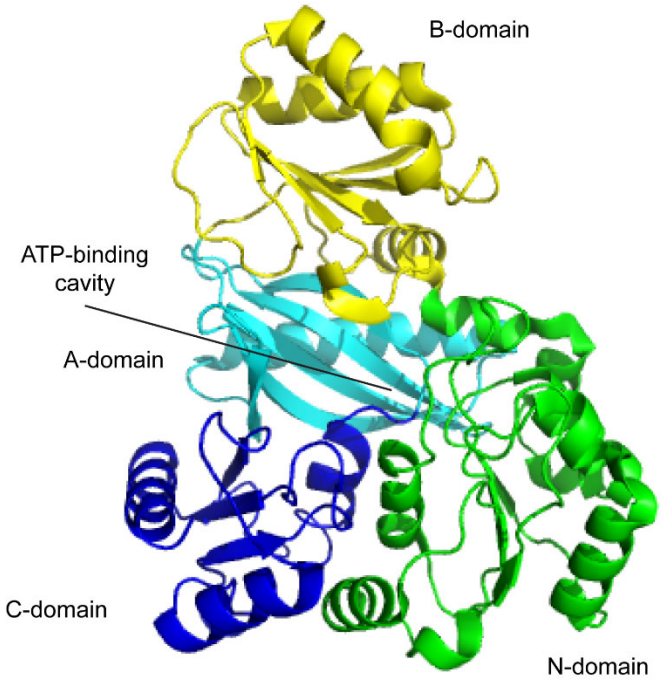


Figure 1.6. Crystal structure of PGM1 (PDB 3WVQ). The large cleft containing the ATP-grasp binding cavity is labeled.

Other ATP-grasp enzymes accept single amino acids, including BacD, BL00235, and RizA. The enzymes BacD and BL00235 are structurally similar, but only share ~20% sequence identity (Fig. 1.7A).[29-30] The enzyme BacD, previously named YwfE, links L-Ala and anticapsin to form bacilysin, an antibiotic and antifungal dipeptide produced by *Bacillus subtilis*.[31-33] Surprisingly, L-Gln is capable of being linked to L-Ala by BacD because anticapsin can act as an analog of L-Gln *in vivo*.[34] However, L-Ala must be the N-terminal residue, thus forming Ala-Gln, and no Gln-Ala.[31] The dipeptide Ala-Ala was also formed by BacD, but no tripeptides were observed by LC-MS.[31] Unlike BacD, BL00235 has a narrow substrate specificity where the N-terminal substrate is limited to L-Met and L-Leu and the C-terminal substrate is limited to L-Ala, L-Ser, L-Thr, and L-Cys.[30] This difference is interesting considering the structural similarity between the two enzymes, with a root-mean-square

difference (r.m.s.d.) between 372 corresponding C^{α} atoms in the enzymes of only 2.71 Å.[30] BL00235 has a deep substrate-binding pocket, and comparison with BacD suggests the dipeptide-binding site is located on top of this pocket.[30] Further comparison with BacD revealed that critical residues for enzymatic activity of BL00235 include Arg290 and Gly293, and necessary residues for recognition of the N-terminal substrate amine include Glu325 and His271.[30] Conversely, Ser294 is necessary for recognition of the C-terminal carboxyl in BL00235, but this is not conserved in BacD, which requires Gly14, Tyr238, Val296, Ile340, Tyr386, and Phe389, and thus forming a hydrophobic cavity.[30] The N-terminal preference of BL00235 is due to this hydrophobic cavity, accepting L-Met and L-Leu as substrates, and the N-terminal preference of BacD is due to the presence of bulky hydrophobic sidechains in the hydrophobic cavity, allowing only small residues such as Gly and L-Ala to enter.[30] Substrate specificity for the C-terminal amino acid is likely controlled by Phe83, Pro85, and Phe156 of BL00235, accepting L-Ala and L-Ser as substrates, and by Ala183 of BacD.[30]

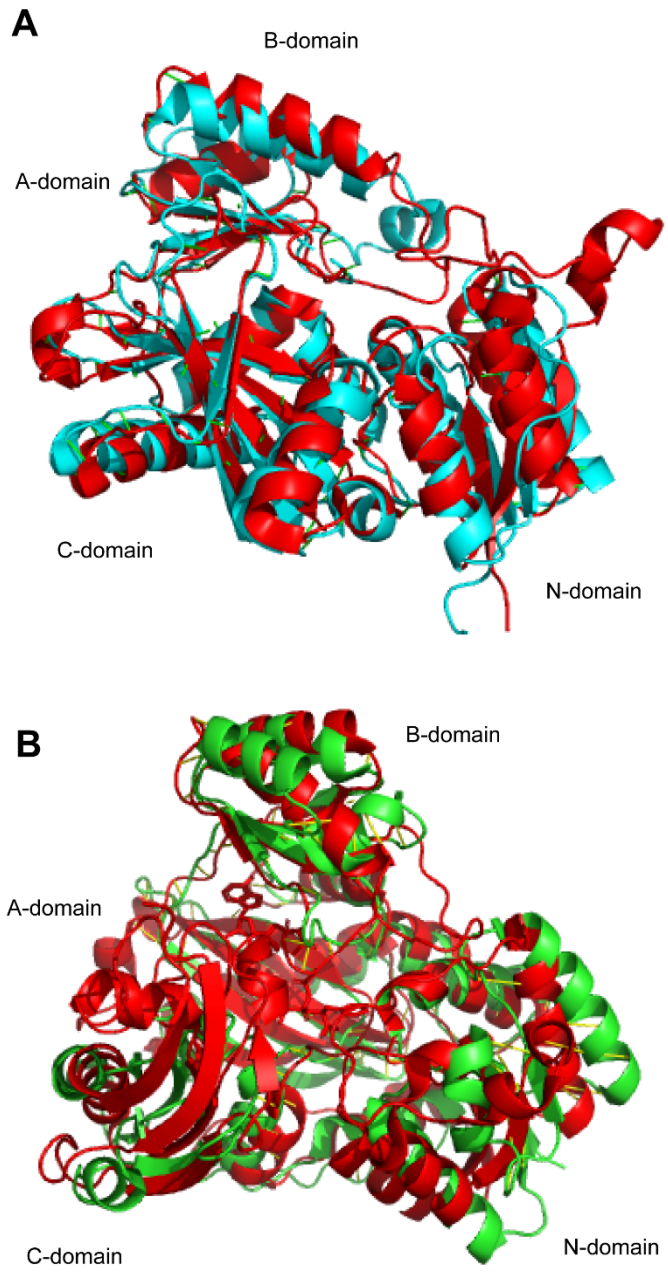


Figure 1.7. Crystal structures of BacD, BL00235, and RizA. (*A*) Alignment of the crystal structures of BacD (red; PDB 3VMM) and BL00235 (blue; PDB 3VOT) shows structural similarity between the enzymes. (*B*) Alignment of the crystal structures of BacD (red) and RizA (green; PDB 4WD3) shows structural similarity between the enzymes.

Similar to both BacD and BL00235 is RizA, which is an ATP-grasp enzyme involved in synthesizing the antifungal linear compound rhizocticin A in *Bacillus subtilis* and has an

overall similar structure to those of both BL00235 and BacD (Fig. 1.7B).[32, 35-37] The respective domains are composed of amino acids 1-121 (N), 122-195 (B), and 196-413 (A+C) in RizA and are essentially identical in BacD and BL00235.[37] However, superposition of RizA with BacD and BL00235 showed poor alignment between large regions of α-helices and β-strands of all three domains.[37] Over 1368 atoms, r.m.s.d. between BacD and BL00235 was 2.195 Å, but r.m.s.d. values between RizA and BacD and between RizA and BL00235 were 5.274 Å and 6.124 Å, respectively.[37] Speculation led to the idea that these larger r.m.s.d. values result from the lack of ADP binding in the crystal structure of RizA.[37] The structure of BacD revealed Lys138, Lys178, Glu226, Glu311, and Glu324 interact with ADP by hydrogen bonding, and these residues are conserved in BL00235 and RizA as well, therefore the ATP-binding pocket is conserved in all three of these L-amino acid ligases.[37] Substrate recognition by BacD is due to Leu12, Trp76, Leu110, Glu273, Glu311, Glu324, and Arg328, and the latter three residues are bound to the peptide bond of the substrate and are conserved in BL00235 and RizA.[37] However, Leu12, Leu110, and Glu273 recognize the sidechains of the dipeptide substrate and are not conserved in BL00235 or RizA.[37]

Since BacD, BL00235, and RizA have been shown to exhibit different substrate specificities, residues conferring substrate-binding pockets for each enzyme likely vary.[37] RizA shows high N-terminal substrate specificity towards L-Arg, linking it to various amino acids.[36] The RizA putative substrate binding pocket is a hydrophobic cavity able to accommodate a large amino acid such as Arg, and the hydrophobic residues likely interact with the hydrocarbon portion of the sidechain of Arg.[37] These reported crystal structures represent a critical starting point for understanding the rationale behind ATP-grasp substrate specificity.

1.4 FUTURE PROSPECTS

More research is needed to fully analyze substrate scope of these enzymes. If enough ATP-grasp enzymes are analyzed to characterize their accepted substrates and crystallized to identify residues necessary for substrate binding, a database can be made. From there, a prediction algorithm can potentially be generated to suggest possible chemistry for each ATP-grasp enzyme by aiding substrate predictions, similar to existing prediction software used for NRPS enzymes.

ATP-grasp enzymes can be used as a search query for genome mining to identify novel natural products. Many NRPS-synthesized natural products have been elucidated through genome mining efforts, but not many ATP-grasp synthesized natural products have. The ATP-grasp enzyme FtyB from *Pseudoalteromonas tunicata* catalyzes the condensation of L-Thr and 3-formyl- L-Tyr to form the natural product 3-formyltyrosine-threonine.[38] This molecule was discovered through homology searching using the dapdiamide ATP-grasp enzyme DdaF as a query, but bioactivity of this molecule has not yet been reported.[38] In depth studies of FtyB substrate scope have also not been reported, so it is unclear whether other substrates could be accepted by this enzyme. This genome mining method would likely lead to novel natural products with potential bioactivity.

Finally, this class of enzymes presents a new method for drug development or synthesis. The ability of these enzymes to catalyze amide bonds could lead to faster synthetic routes than total synthesis and may be simpler to use than solid-phase peptide synthesis. In particular, TabS could be used to generate many different dipeptides due to its lax substrate specificity. Overall, this class of enzymes holds promise for identifying novel natural products.

CHAPTER 2: PHEVAMINE A AND MANGOTOXIN: SMALL MOLECULES PRODUCED BY *PSEUDOMONAS SYRINGAE* THAT AFFECT PLANT INFECTIONS

The work on phevamine A was completed in collaboration with Tatiana Mucyn and Jeff Dangl at UNC. They created genetic knockouts of the *hsv* cluster and performed all plant assays with *Pto* and the phevamines. Phylogenetic analysis of *hsv* was conducted by Omri Finkel. Chemical synthesis of phevamine A was conducted by Jon Patteson and Elisabetta Massolo.

2.1 INTRODUCTION

Bacterial small molecules play key roles in pathogen–plant interactions. *Pseudomonas syringae*, a phylogenetically diverse species of bacteria responsible for many crop diseases, deploys a myriad of virulence factors, including secreted protein effectors delivered to host cells by the type III secretion system (T3SS) and small molecules.[39-40] Coronatine, syringomycin, syringopeptin, tabtoxin, and phaseolotoxin are among the small molecule virulence factors produced by various strains of *P. syringae*, acting as phytohormone mimics or phytotoxins.[39] For example, coronatine disrupts plant immune signaling by mimicking the structure of the plant hormone jasmonic acid-isoleucine,[39, 41] and phaseolotoxin contributes to virulence by interfering with the synthesis of L-arginine and polyamines.[42-43] While many protein effectors have been extensively studied in *P. syringae*, only a limited number of small molecules have been identified. The genomes of *P. syringae* and other phytopathogens encode potential biosynthetic enzymes.[44-45] Identifying the cryptic small molecules synthesized by these enzymes may reveal new virulence mechanisms that could be targeted to control plant infection.

Many phytopathogenic virulence factors function to suppress plant immune responses.[46-47] The first layer of the plant immune system relies on the recognition of conserved microbe-associated molecular patterns (MAMPs) by host pattern recognition receptors to generate MAMP-triggered immunity (MTI).[48] The MTI signaling cascade includes receptor kinase activation, reactive oxygen species (ROS) production, calcium influx, mitogen-activated protein kinase (MAPK) activation, defense gene activation, and callose deposition at the plant cell wall (Fig. S2.1); each of these steps can be targeted by virulence factors.[48] The sigma factor HrpL controls the expression of *P. syringae* virulence factors including the T3SS and its associated effector genes.[49] HrpL also regulates genes responsible for the biosynthesis of coronatine in some, but not all, *P. syringae* strains.[13, 50] Our recent study identified a number of uncharacterized HrpL-regulated genes that encode putative biosynthetic enzymes for small molecules.[13] We therefore proposed that these HrpL-regulated genes are responsible for synthesizing small molecules that play a role in virulence.[13]

In this chapter, we focus on two biosynthetic operons: the HrpL-regulated three-gene cluster *PSPTO_0873–0875* from *P. syringae* pv. *tomato* DC3000 (*Pto*),[13, 51] and the six-gene cluster *mboA–F* from *P. syringae* pv. *syringae* UMAF0158 (UMAF).[14] The *Pto* operon shares 66-85% homology with *hrp* (hypersensitive response and pathogenicity)-associated systemic virulence (*hsv*) genes from *Erwinia amylovora.*[52] The *Pto hsv* operon was implicated in the virulence of *E. amylovora* on apple shoots,[53] but the small molecule(s) produced by *hsv*-encoded enzymes were unknown. The mangotoxin biosynthetic operon, *mboA–F*, shares 26–27% sequence identity with the proteins encoded by the *Pto hsv* cluster. Mangotoxin, synthesized by the enzymes encoded in the *mbo* cluster, is an antimetabolite that inhibits plant ornithine N-acetyl transferase, an essential enzyme for ornithine and arginine biosynthesis.[15]

The bioactivity of mangotoxin was reported in 2003 and the *mbo* cluster was identified in 2012, however, the structure of mangotoxin still remains unknown.[14-15]

Here, we confirm that the *Pto hsv* operon acts as a *P. syringae* virulence factor on *Arabidopsis thaliana* and demonstrate that the *hsv* operon dampens host immune responses in both *A. thaliana* and *Nicotiana benthamiana.* We demonstrate that the enzymes encoded by the *hsv* operon synthesize a small molecule that is important for virulence. By integrating heterologous expression, metabolomics, and in vitro biosynthesis, we identify a novel bioactive small molecule, phevamine A, the conjugate of L-phenylalanine, L-valine and a modified spermidine. We further show that phevamine A suppresses the potentiation of MAMP-induced ROS bursts by spermidine and L-arginine. Thus, phevamine A is a new small molecule virulence factor that promotes bacterial growth and virulence, in part by suppressing plant immune response. We also show that the *hsv* and *mbo* enzymes share fairly high (~27%) sequence identity, but they prefer different substrates and therefore likely synthesize different small molecules.

2.2 RESULTS AND DISCUSSION

2.2.1 The P. syringae hsv operon promotes virulence and suppresses defense responses.

To investigate the function of the *hsv* operon, we first generated independent clean deletion mutants in *Pto*. The parental strain, *Pto*, and the two mutants, *Pto_*Δ*hsv1*, *Pto_*Δ*hsv2* were used to infect *A. thaliana* Col-0 seedlings by dip inoculation and monitor bacterial growth. We observed that the mutants grew slightly less than the *Pto* strains (Fig. S2.2) in two independent experiments, but this reduced growth trend was not statistically significant. This observation is not surprising, as *Pto* is an aggressive pathogen on *A. thaliana* and its large suite

of virulence factors can act collectively.[54] It is therefore rare to observe significant loss of virulence phenotypes when deleting a single candidate virulence factor from this strain. To circumvent this problem, we used a weakly pathogenic *Pto* derivative deficient in coronatine production (*Pto* DC3118, hereafter *Pto*-Cor $^{-}$)[54] to generate independent *hsv* mutants *Pto*-Cor $^{-}$ Δ*hsv1*, and *Pto*-Cor $^{-}$ Δ*hsv2*. Three days post-infection, significantly less bacteria for *Pto*-Cor $^{-}$*Δhsv*1 and *Pto*-Cor $^{-}$*Δhsv*2 were recovered from the seedlings than *Pto*-Cor $^{-}$ (Fig. 2.1A). This result indicates that *hsv* is required for full virulence *in planta*.

We next addressed whether the *hsv* operon had an effect on host callose deposition, a cell wall reinforcing response typically triggered during MTI (Fig. S2.1).[55-56] Several type III secretion effector proteins can inhibit callose deposition[57-61] and thus may obscure phenotypes resulting from *hsv* deletion. We therefore generated two independent *hsv* clean deletions in *Pto*D28E (*Pto*D28EΔ*hsv*1 and *Pto*D28EΔ*hsv*2), a strain lacking 28 different type III secretion effectors.[62] *Pto*D28E, *Pto*D28EΔ*hsv*1 and *Pto*D28EΔ*hsv*2 were used to infiltrate four-week-old leaves of *A. thaliana*. Both *Pto*D28EΔ*hsv* alleles induced higher callose deposition than *Pto*D28E (Fig. 2.1B); complementation of *Pto*D28EΔ*hsv*2 with constitutively expressed *hsv* reduced the level of callose back to that induced by *Pto*D28E (Fig. 2.1C). These results demonstrate that *hsv* suppresses a late marker of plant defense at the cell wall.

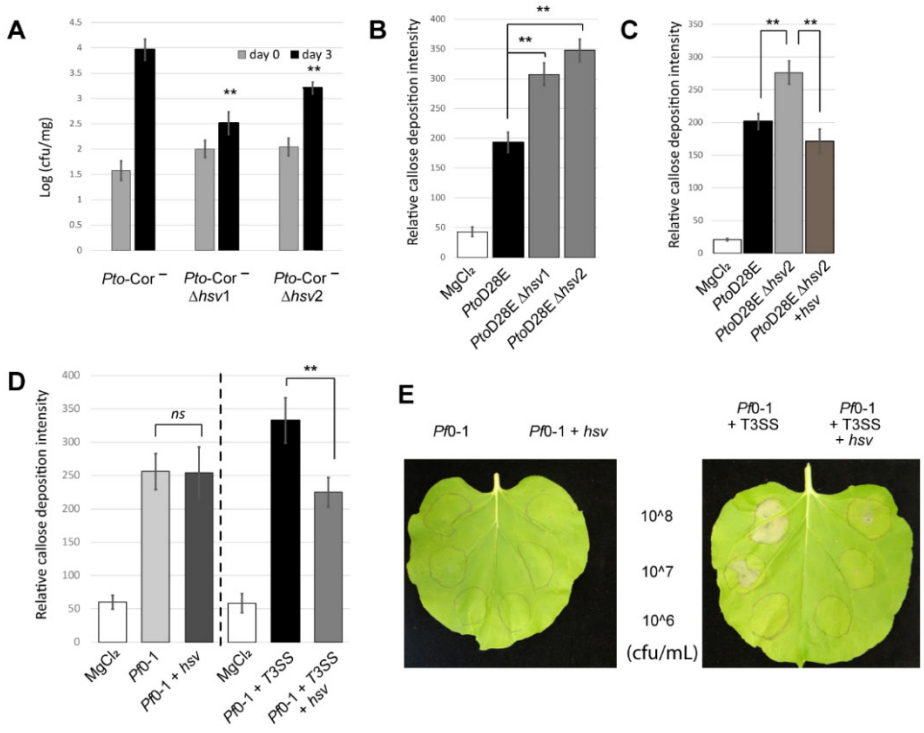


Figure 2.1. The hsv operon is required for virulence and suppresses defense responses. (*A*) *Pto*-Cor⁻Δ*hsv* mutants display reduced growth on *A. thaliana* seedlings compared to *Pto*-Cor⁻. This experiment was repeated three times with similar results. (*B*) *Pto*D28EΔ*hsv* mutants induce higher callose deposition than *Pto*D28E on 4-week-old *A. thaliana* leaves. The experiment is a representative of 4 independent replicates. (*C*) Complementation of *Pto*D28EΔ*hsv2* in *Pto*D28EΔ*hsv2* + *NptII::hsv* re-establishes callose deposition to levels triggered by *Pto*D28E. This assay is a representative of two independent replicates. (*D*) The *hsv* operon suppresses callose deposition when co-expressed in *Pf*0-1 with T3SS, but not when expressed in *Pf*0-1. The callose deposition was monitored as in (*B*) after infiltration of *Pf*0-1 or *Pf*0-1 expressing *hsv*. This assay was performed twice with similar results. *Pf*0-1+T3SS and *Pf*0-1+T3SS expressing *hsv* were tested separately with at least three replicate experiments for each strain. (*E*) The *hsv* operon in *Pf*0-1 suppresses T3SS-mediated cell death in *N. benthamiana*. *Pf*0-1 derived strains were inoculated at three different concentrations (cfu/mL). Pictures were taken 20 hours post-infiltration. This assay was repeated 4 times with similar results. ** indicates *t*-test, *p* value < 0.01. Error bars represent ± standard error. Cfu/mg: colony forming units per milligrams of plant tissue. Relative callose intensity represents the number of callose deposits observed per field.

Conversely, we also examined whether *hsv* expression is sufficient to suppress plant callose deposition induced by a non-pathogenic bacterium. The *hsv* operon was constitutively expressed in *Pseudomonas fluorescens Pf*0-1 (*Pf*0-1), a strain that lacks both *hsv* and a type III secretion system (T3SS), but still induces callose deposition (Fig. 2.1D).[63] Expression of *hsv* had no effect on the level of *Pf*0-1-induced callose deposition (Fig. 2.1D). A *Pf*0-1 derivative engineered to contain a complete *P. syringae* T3SS locus (*Pf*0-1+T3SS)[64] induced more callose than wildtype *Pf*0-1 (Fig. S2.3), suggesting that *Pf*0-1 T3SS components are recognized by the plant.[65] Expression of *hsv* in the *Pf*0-1+T3SS strain reduced the level of callose back to the level elicited by wildtype *Pf*0-1 (Fig. 2.1D, Fig. S2.3). Thus, *hsv* suppresses the callose deposition induced by *Pf*0-1 T3SS components. To extend this observation, we evaluated the activities of these *P. fluorescens* strains on a second plant species, *Nicotiana benthamiana.* Infiltration of *N. benthamiana* leaves with *Pf*0-1 or *Pf*0-1 expressing *hsv* had no effect; the tissue infiltrated with bacteria was identical to the non-infiltrated tissue (Fig. 2.1E). In contrast, infiltration with *Pf*0-1+T3SS elicited a dose-dependent leaf cell death. This response was suppressed by *hsv* expression (Fig. 2.1E). Thus, *hsv* suppresses the cell death caused by a component of the T3SS or the linked harpin gene transferred from *P. syringae* pv. *syringae* to *Pf*0-1.[66-67]

2.2.2 hsv-encoded enzymes synthesize three related small molecules.

The *hsv* operon encodes three enzymes, including a putative amidinotransferase HsvA and two putative ATP-grasp type enzymes HsvB and HsvC (Fig. 2.2A). To test the hypothesis that these *hsv*-encoded enzymes synthesize small molecules, we conducted *hsv* heterologous expression and comparative metabolomics experiments. The *hsv* cluster was overexpressed in

E. coli as a heterologous host to enhance small molecule production. Organic extracts of culture supernatants were analyzed using liquid chromatography-coupled high-resolution mass spectrometry (LC-HRMS) to generate metabolomic profiles for comparison. Three species with mass-to-charge ratios (*m/z*) of 287.255, 434.324, and 581.392 were present at high levels in the *hsv*-expressing *E. coli,* but were absent in the control *E. coli* carrying the empty vector (Fig. 2.2B, Fig. S2.4–2.6). The species with the *m/*z of 287.255 and 434.324 were also detected by LC-HRMS in the culture extract of the wildtype *Pto*, but not in *Pto*Δ*hsv* (Fig. 2.2C, Fig. S2.5). The production of the metabolite with the *m/z* of 434.324 is also observed in *Pto* overexpressing *hrpL,* but is reduced by *hrpL* deletion (Fig. 2.2C), consistent with the regulation of *hsv* by *hrpL.*[13] This metabolite was detected at the same level in the supernatants of both *Pf*0-1 and *Pf*0-1+T3SS expressing *hsv* (Fig. S2.7), indicating that secretion of the compound is independent of T3SS.

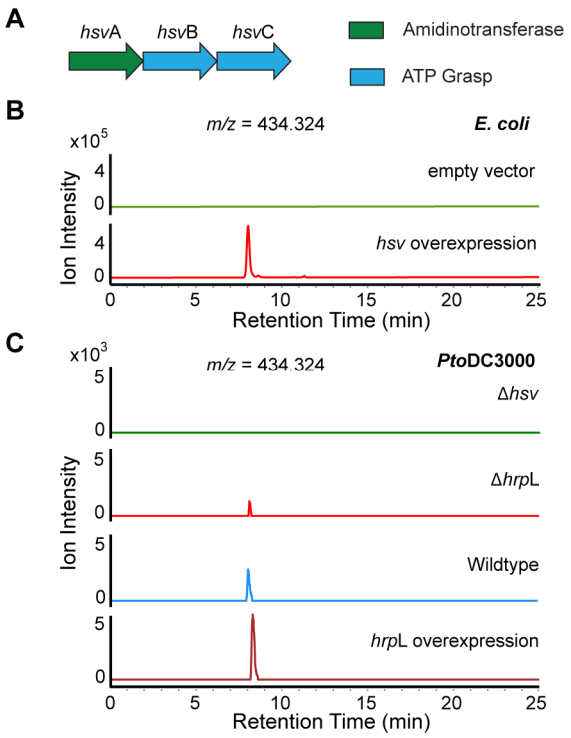


Figure 2.2. Metabolomics reveal the small molecules synthesized by hsv-encoded enzymes. (*A*) Graphical representation of the *hsv* operon. (*B*) Comparative metabolomics analysis in *E. coli*. Overexpression of *hsv* produces a new peak with a mass-to-charge (*m/z*) ratio of 434.324 (green), which is absent in *E. coli* harboring the empty vector (red). (*C*) Comparative metabolomics analysis in *Pto*. The molecule with the *m/z* of 434.324 is absent in Δ*hsv* (green) and reduced in Δ*hrpL* (red), but present in both wildtype *Pto* (blue) and *Pto* overexpressing *hrpL* (maroon). Extracted ion chromatogram for 434.324 is shown in (*B*) and (*C*). The slight change in retention time is due to the hydrophilic nature of this molecule. Experiments presented in (*B*) and (*C*) were repeated at least 3 times.

We analyzed the structures of the identified metabolites from *E. coli* and *P. syringae* strains expressing the *hsv* operon by tandem MS, which revealed that these metabolites share similar MS fragments (Fig. S2.8, Fig. S2.9) and likely possess related structures. The mass differences between 287.255 and 434.324, and between 434.324 and 581.392 are both 147.068, corresponding to the mass of a phenylalanine in an amide linkage. Based on these data, we

propose that the structures of the compounds with the *m/z* of 434.324 and 581.392 contain one and two phenylalanines, respectively. We name the identified metabolites phevamine A (*m/z* 434.324; PHVA), phevamine B (*m/z* 581.392; PHVB) and pre-phevamine (*m/z* 287.255; pPHV).

2.2.3 In vitro biosynthesis of phevamines reveals novel structures.

To characterize the structures of phevamines, we reconstituted the biosynthetic enzymes encoded by the *hsv* operon in vitro. HsvA shares homology with amidinotransferases that transfer an amidino group from arginine to an amine, suggesting the presence of an amidino group. HsvB and HsvC belong to the ATP-grasp enzyme superfamily consisting of ATP-dependent enzymes that condense carboxylic acids with amines, suggesting the presence of amide linkages in the phevamines and pre-phevamine.[1] HsvA, HsvB, and HsvC were overexpressed in *E. coli* and purified as recombinant proteins. We reconstituted the activity of HsvA by incubating L-arginine, the donor of the amidino group for amidinotransferases, in the presence of a variety of physiologically relevant amines as acceptors, including L-lysine as well as spermidine and spermine, which are polyamines produced by *P. syringae.*[68-69] We found that HsvA preferentially modifies spermidine and generates amidinospermidine (Fig. 2.3A, Fig. S2.10, Fig. S2.11). Subtracting the mass of amidinospermidine from that of pre-phevamine results in a mass consistent with a valine in an amide linkage. We therefore proposed amidinospermidine is condensed with L-Val as the second biosynthetic step to generate pre-phevamine (Fig. 2.3A). Both ATP-grasp enzymes, HsvB and HsvC, were examined for this condensation activity in the presence of amidinospermidine, L-Val, ATP and Mg^{2+}. Only HsvC catalyzes pre-phevamine formation (Fig. 2.3A, Fig. S2.10). Data from LC-

HRMS analyses and phosphate release assays suggest that L-phenylalanine is a structural component of the phevamines (Fig S2.12). To validate this proposal, we incubated pre-phevamine with HsvB and L-Phe, and the formation of phevamine A and phevamine B was indeed observed (Fig. 2.3A, Fig. S2.10, Fig. S2.13). This result also indicates that HsvB can catalyze two rounds of condensation with L-Phe. Tandem MS analysis of phevamine A suggests that the amidino group is connected to the propylamine end of spermidine (Fig. 2.3B, Fig. S2.9).

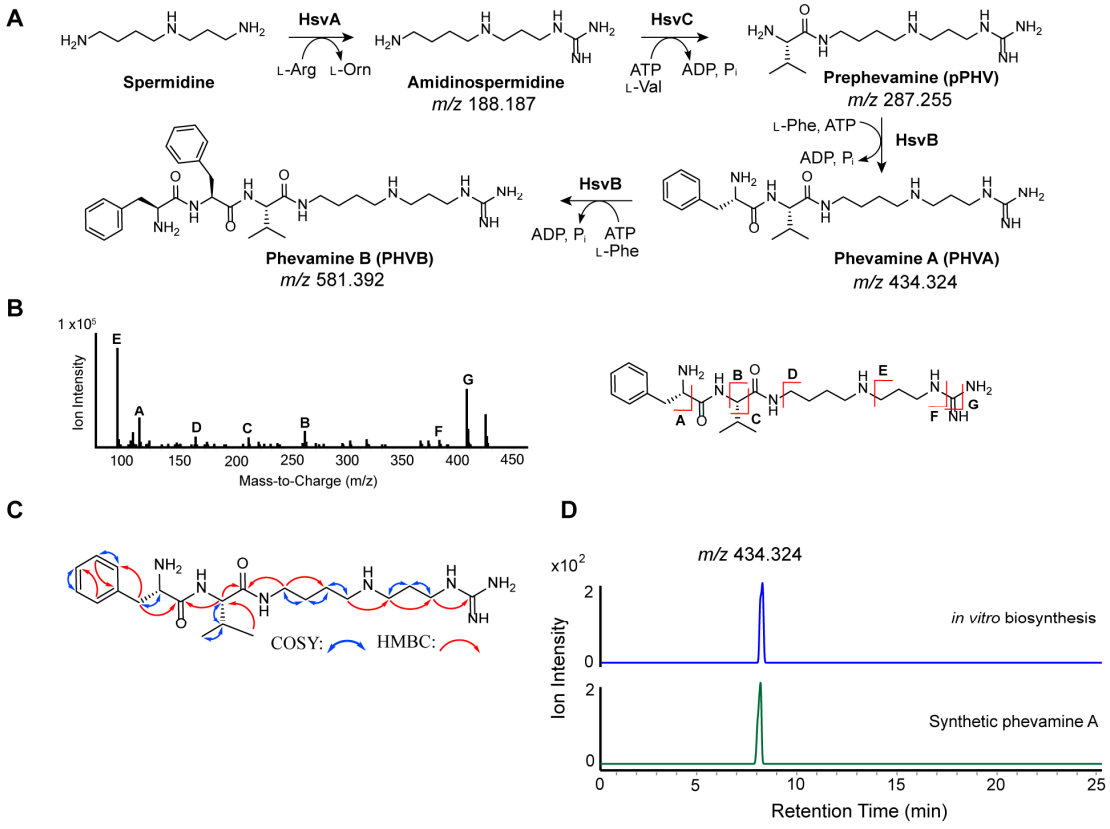


Figure 2.3. Structural characterization and biosynthesis of pre-phevamine and phevamines. (*A*) Biosynthetic pathway of pre-phevamine and the phevamines based on in vitro characterization. (*B*) Structural analysis of phevamine A produced by *E. coli* overexpressing *hsv* using tandem MS. *(C)* (^{1}H,^{13}C)-HMBC correlations of in vitro synthesized and purified phevamine A. (*D*) Extracted ion chromatogram of phevamine A (*m/z*, 434.324) produced via in vitro enzymatic synthesis (blue) or total chemical synthesis (green).

Having identified that phevamines consist of L-Phe, L-Val, and amidinospermidine, we next determined the connectivity of these components by NMR analysis. Due to high polarity, phevamines were difficult to isolate from bacterial culture extracts. To isolate sufficient materials for NMR, we developed in vitro biosynthetic methods for phevamine A and phevamine B. From 16 mL of in vitro enzymatic assays, 5.7 mg of phevamine A and 5.5 mg of phevamine B were purified. These compounds were analyzed by ^{1}H, (^{1}H,^{1}H)-COSY, (^{1}H,^{13}C)-HMBC, and (^{1}H,^{13}C)-HSQC NMR experiments (Fig. 2.3C and Fig. S2.14–2.21). The NMR data supports that the amidino group is linked to the propylamine side of spermidine and that phevamine B contains an additional L-Phe at the *N*-terminus of phevamine A. Phevamine B was only detected under conditions of heterologous expression and in vitro synthesis, but not in the culture extract of *Pto* (Fig. S2.6), suggesting that phevamine B is unlikely a physiologically relevant molecule. The production of phevamine B may be due to the overexpression of HsvB in *E. coli* and the high concentration of HsvB and L-Phe added in vitro. Thus, we focused on phevamine A for further structural characterization. To confirm the relative stereochemistry, we developed a total synthesis method for phevamine A. Synthetic and in vitro isolated phevamine A exhibit identical ^{1}H and ^{13}C NMR signals, and the same LC retention time and MS fragmentation pattern as the species detected in the bacterial culture extract (Fig. 2.3D, Fig. S2.14–2.22). These results confirmed the structure assignment for phevamine A and provided materials for biological activity testing of this molecule. In an initial activity test, we examined the ability of phevamine A to bind ferric iron using a chromeazurol S assay and observed no significant binding to ferric iron (Fig. S2.23), suggesting phevamine A is unlikely acting as a siderophore.

2.2.4 Comparative genomics suggests a potential host target for phevamines.

We investigated the phylogenetic distribution of *hsv* in bacterial genomes using MultiGeneBlast that identifies homologous gene clusters based on sequence similarity and gene synteny.[70] We found that *hsv* is widely distributed across plant-pathogenic bacterial genera including *Pseudomonas*, *Erwinia*, and *Pantoea*. We compared the distribution of the *hsv* operon to those biosynthetic gene clusters encoding known *P. syringae* phytotoxins, focusing on coronatine, mangotoxin, phaseolotoxin, tabtoxin, and syringolin (Fig. 2.4A; Table S2.1).[39, 71] The *hsv* operon is present in ~36% of *P. syringae* (107 out of 292) and is the most widely-distributed gene cluster of the six analyzed. The *hsv* operon rarely co-occurs with any of the aforementioned small molecule biosynthetic clusters. This anti-correlation is especially prominent between *hsv* and phaseolotoxin biosynthetic genes in the genomes of otherwise extremely closely related strains of *P. syringae* pv. *actinidiae.*[72] We previously showed that functionally redundant virulence factors rarely co-occur in the same strain.[73] Specifically, the small molecule coronatine and three different type III secretion system protein effectors use four distinct mechanisms to target the same host defense signaling pathway, and only genes responsible for a single mechanism are typically found in any given *P. syringae* genome.[73] Based on the clear anti-correlation of *hsv* with the phaseolotoxin biosynthetic genes, and the knowledge that phaseolotoxin targets host arginine and polyamine biosynthesis by inhibiting ornithine carbamoyltransferase and ornithine decarboxylase (Fig. 2.4B),[42-43, 74] we hypothesized that the action of phevamine A involves polyamines and/or arginine. This hypothesis is consistent with previous implication of polyamine function in plant immune responses.[75-76] Furthermore, our genomic analysis revealed that many *P. syringae* strains do not contain the biosynthetic gene clusters for phevamine A or the five other phytotoxins

described herein (Fig. 2.4A). These strains could produce different small molecules that might play a role in bacteria–plant interaction.

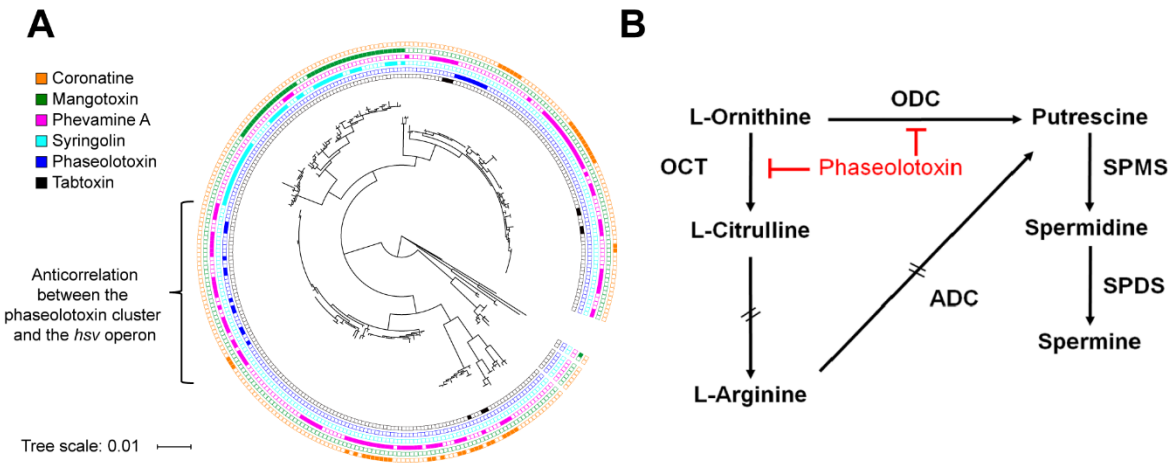


Figure 2.4. Distribution of the hsv operon and other toxins among P. syringae and graphical representation of polyamine biosynthesis and action of phaseolotoxin. (*A*) Phylogenetic tree of 292 *P. syringae* strains and the distribution of coronatine, mangotoxin, phevamine A, syringolin, phaseolotoxin, and tabtoxin biosynthetic genes, from the outer ring to the inner ring. The tree scale represents the nucleotide substitution per site. (*B*) Phaseolotoxin inhibits the arginine and polyamine pathways by suppressing the activity of ornithine decarboxylase (ODC) and ornithine carbamoyltransferase (OCT). ADC: arginine decarboxylase, SPDS: spermidine synthase, SPMS: spermine synthase. Double dashed arrows represent pathways with intermediates not shown. Adapted from [77].

2.2.5 Polyamines and arginine potentiate an early MTI response.

We investigated the effect of arginine and polyamines on the MTI response induced by the well-studied MAMP flg22, a short peptide derived from the *Pseudomonas* flagellin.[55] We monitored the ROS burst following recognition of flg22 by the pattern recognition receptor FLS2 (Fig. S2.1).[78-79] The polyamines putrescine, spermidine, and spermine potentiated the flg22-induced ROS burst in *N. benthamiana,* resulting in an earlier and larger amplitude response (Fig. 2.5A, Fig. 2.5B, Fig. S2.24A). The potentiation was dose-dependent (Fig.

S2.24B). Spermidine alone induced only a slight ROS burst (Fig. S2.24C),[76] and the increase of the flg22-induced ROS burst in the presence of spermidine was higher than an additive effect (Fig. S2.24C). The potentiation by spermidine was also observed in *A. thaliana*, but was not observed in the flg22 receptor mutant *fls2* (Fig. 2.5C, Fig. S2.25A, Fig. S2.25B). Spermidine potentiation was also observed using the elicitor elf18 (Fig. S2.25C).[80] Surprisingly, the Arabidopsis Ca^{++} burst induced by flg22 was reduced in the presence of spermidine (Fig. 2.5D, Fig. S2.1, Fig. S2.25D). Arginine had a similar effect on this early MTI response, while L-citrulline did not (Fig. S2.26, Fig. S2.27). Thus, spermidine and arginine potentiate MAMP-induced ROS bursts in both plant species, and spermidine modulates the flg22-induced Ca^{++} burst, at least in *A. thaliana*. These results are consistent with previously proposed functions of polyamines in plant defense responses,[75-76] including upregulation of polyamine biosynthesis in plants following challenge with bacterial pathogens,[81-84] subsequent polyamine transport to the apoplast,[81] and increased resistance toward *Pseudomonas viridiflava* observed after exogenous addition of spermine.[84]

2.2.6 Phevamine A suppresses spermidine and arginine potentiation of the flg22-induced ROS burst.

The anti-correlation of phevamine A and phaseolotoxin production suggested that phevamines might impair the effect of spermidine and arginine on the early MAMP-dependent ROS burst. We therefore tested whether phevamine A could impair the effect of exogenously supplied spermidine and arginine in the MTI-induced ROS burst potentiation assay. Leaf disks of *N. benthamiana* were challenged with flg22 in the presence of phevamine A, spermidine, or both (Fig 2.5A). We observed that phevamine A suppressed the spermidine potentiation of the flg22-induced ROS burst in *N. benthamiana*, but did not directly affect the flg22 response (Fig.

2.5A). We noted that pre-phevamine also suppressed the spermidine potentiation of the flg22-induced ROS burst, but phevamine B did not (Fig. 2.5B). This result is consistent with our observation that phevamine B is not naturally produced by *Pto* (Fig. S2.6). In contrast, only phevamine A suppressed the spermidine-mediated flg22-induced ROS burst in *A. thaliana* leaf disks in a similar experimental set up (Fig. 2.5C). This indicates that the molecular mechanism(s) of the spermidine potentiation might differ slightly between *N. benthamiana* and *A. thaliana*.

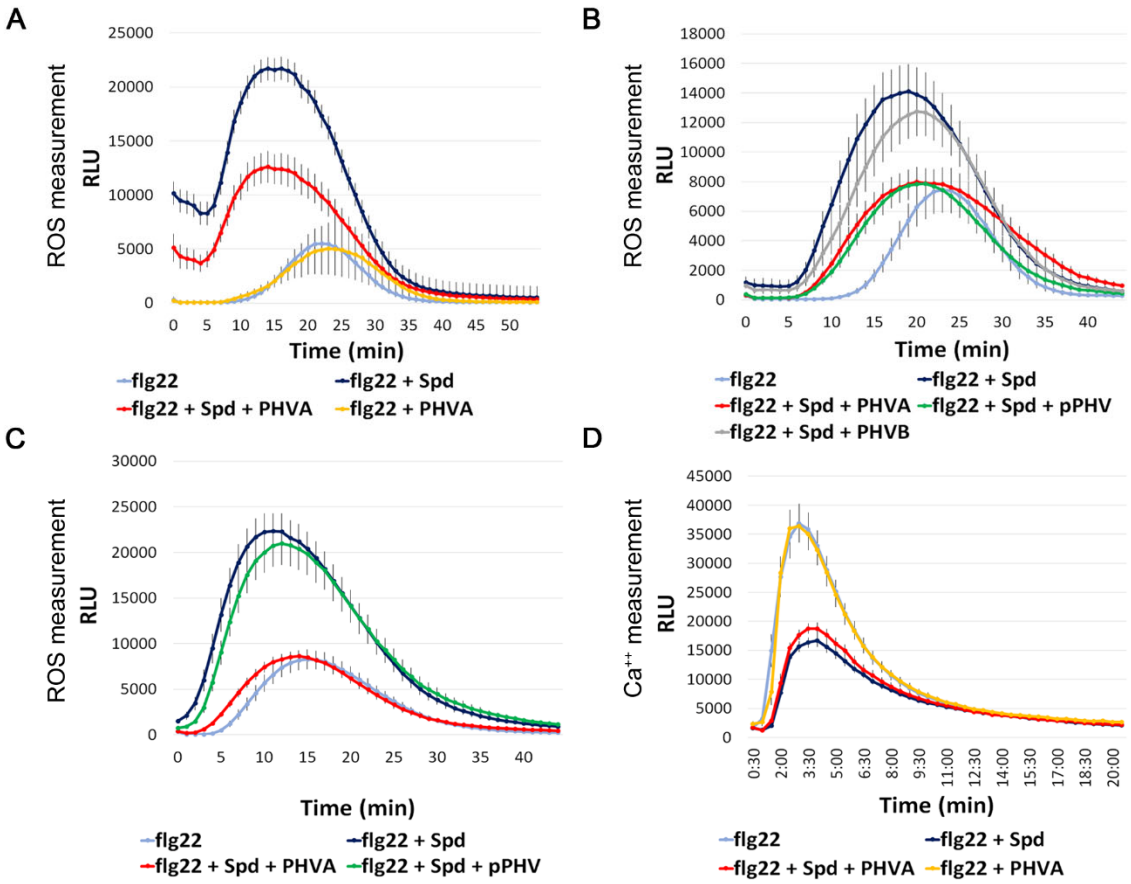


Figure 2.5. Effect of the phevamines and pre-phevamine on the spermidine potentiation of the ROS burst and suppression of the calcium burst. (*A*) Phevamine A suppresses the spermidine potentiation of the flg22-induced ROS burst in *N. benthamiana*, but does not affect the flg22-induced ROS burst directly. Leaf disks were treated with 50 nM flg22, with or without spermidine at 400 µM, and with or without phevamine A at 400 µM. (*B*) Phevamine

A and pre-phevamine suppress the spermidine potentiation in *N. benthamiana*, but phevamine B does not. Leaf disks were treated with flg22 at 10 nM, spermidine, phevamines and pre-phevamine were used at 300 μM. (*C*) In *Arabidopsis*, phevamine A suppresses the spermidine potentiation, while pre-phevamine does not. Leaf discs were treated with flg22 at 10 nM, spermidine and phevamines were used at 300 μM. (*D*) Phevamine A does not affect the flg22-induced Ca^{++} burst, or the reduction of this burst by spermidine. Leaf disks were treated with flg22 at 20 nM, spermidine and phevamine A were used at 300 μM. Error bars represent standard errors. Relative luminescence units: RLU, phevamine A: PHVA, phevamine B: PHVB, pre-phevamine: pPHV, Spermidine: Spd. All experiments presented here were repeated at least 3 times.

Spermidine affects both the flg22-induced ROS and Ca^{++} bursts (Fig. 2.5D, Fig. S2.24C). We therefore tested the effect of phevamine A on the flg22-induced Ca^{++} burst in the presence and absence of spermidine using a transgenic *A. thaliana* line that expresses the aequorin reporter.[85] As noted above, spermidine suppressed the flg22-induced Ca^{++} burst, but the addition of phevamine A did not affect this suppression. Phevamine A also had no direct effect on the Ca^{++} burst induced by flg22 (Fig. 2.5D). These data suggest that phevamine A acts downstream of the Ca^{++} burst.

Additionally, both phevamine A and pre-phevamine were tested for their effect on the ROS potentiation mediated by arginine in both *N. benthamiana* and *A. thaliana*. Phevamine A, but not pre-phevamine, could suppress the arginine-mediated potentiation of the flg22-induced ROS burst in *N. benthamiana* (Fig. S2.27). This may imply that the potentiation mediated by spermidine and arginine in *N. benthamiana* involves slightly divergent molecular mechanism(s), supported by the different potentiation phenotypes for pre-phevamine. Similarly, only phevamine A and not pre-phevamine could inhibit the arginine potentiation of the flg22-induced ROS burst in *A. thaliana*.

Overall, we observed that phevamine A consistently suppressed both the spermidine- and arginine-mediated potentiation of the flg22-induced ROS burst across two divergent plant species. Thus, phevamine A is likely the most physiologically relevant small molecule produced by the *hsv* operon in *Pto*.

2.2.7 In vitro reconstitution of mbo enzyme activity reveals different substrate preferences from hsv enzymes.

To elucidate the structure of mangotoxin, we purified the biosynthetic enzymes MboB, MboC, MboD, and MboE for in vitro enzyme assays (Fig 2.6). Similar to the *hsv* enzymes, MboC and MboD are putative ATP-grasp enzymes and MboE is a putative amidinotransferase. MboB is likely a (2Fe-2S)-containing ferredoxin oxidoreductase based on sequence homology. We have shown that the amidinotransferase HsvA catalyzes transfer of the amidino group from L-Arg to spermine, spermidine, and L-Lys.[12] MboE was also tested for amidinotransferase activity by combining L-Arg, the typical amidino donor, with L-Lys, L-Cys, L-Ala, Gly, spermine, or spermidine, but no products from the transfer of the amidino group were observed in the presence of these substrates. Therefore, MboE likely performs different reactions or modifies different substrates from HsvA. Also, since spermidine, glycine, and lysine are the only reported amine acceptors for amidinotransferase enzymes, it is unclear what the native substrate of MboE is. Additionally, deletion of *mboE* in UMAF0158 only decreased mangotoxin production slightly and did not diminish the production fully, therefore MboE is most likely unnecessary for the synthesis of the bioactive mangotoxin.[14]

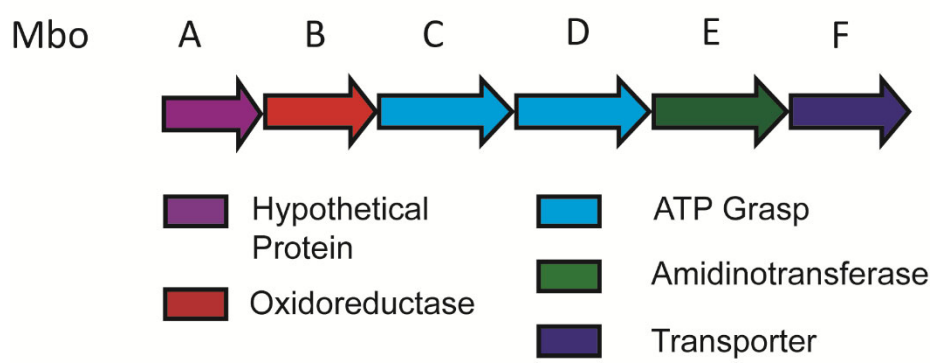


Figure 2.6. Graphical representation of the mbo operon. The *mbo* gene cluster encodes an amidinotransferase (green), two ATP-grasp enzymes (blue), and an oxidoreductase (red).

The *mbo* cluster encodes two ATP-grasp enzymes, MboC and MboD, like the *hsv* cluster, so we predict that mangotoxin also contains two amide bonds. To identify the substrates of MboC and MboD, all proteinogenic amino acids were screened in vitro as potential substrates. Based on LC/MS analysis, MboC catalyzes the formation of the following L-amino acid peptides: Arg-Ala, homoarginine (hArg)-Ala, hArg-Cys, Arg-Ala-Leu/Ile, Arg-Ala-Lys, Arg-Ala-Met, hArg-Ala-Val, hArg-Ala-Leu/Ile, and hArg-Ala-Met (Fig. 2.7A, 2.7B, Fig. 2.8A–C). However, the activity of MboD could not be reconstituted in the presence of the 20 canonical amino acids or spermine, spermidine, ornithine, or homoarginine. This result suggests that MboD acts on a nonproteinogenic amino acid or an amino acid that is modified first by MboB, or a peptide generated by MboC, which can catalyze the formation of several di- and tripeptides from free amino acids. Interestingly, MboC formed three times as much hArg-Ala than Arg-Ala. If MboE could act on L-Lys as a substrate first to form hArg, then MboC could form the hArg-Ala dipeptide, but MboE is not able to catalyze this reaction. Further, the Arg-Ala-Ala dipeptide was observed in the presence of both MboE and MboC, but not in the presence of MboC alone. Therefore, MboE may act as a chaperone for or

complements MboC, consistent with the observation that deleting *mboE* reduced mangotoxin production, but did not completely diminish it.

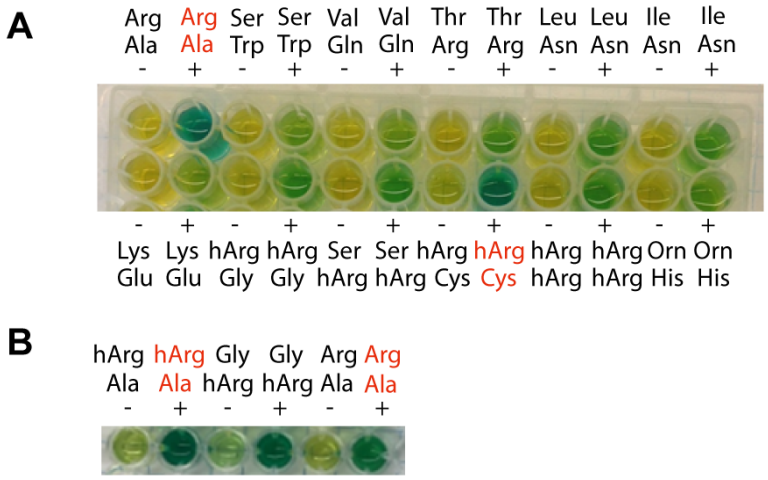


Figure 2.7. Preliminary phosphate release colorimetric assays suggest the substrate scope of MboC. The amino acid pairs listed were incubated with the necessary cofactors without (-) or with (+) MboC. Darker assays show release of free phosphate, suggesting enzyme activity. Assays showing the darkest colors were analyzed by LC-MS to confirm formation of the dipeptide and are highlighted in red.

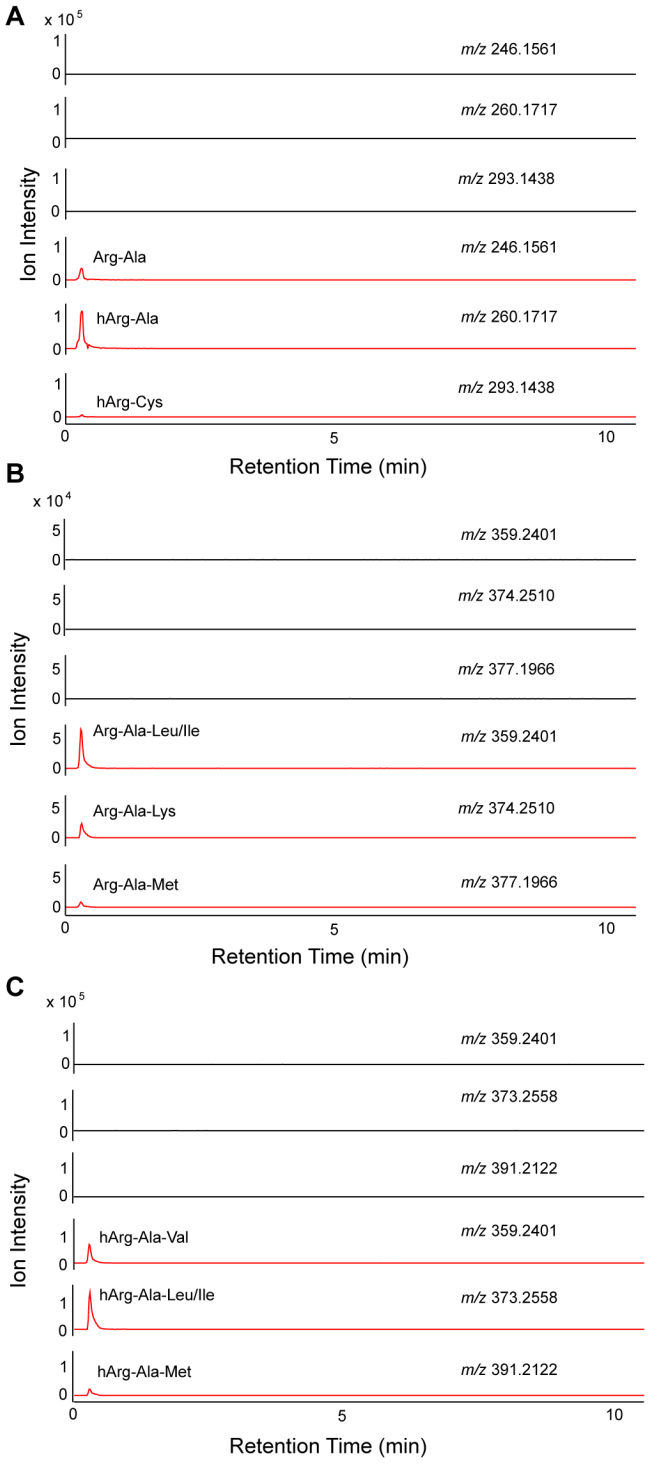


Figure 2.8. MboC catalyzes the formation of di- and tri-peptides in the presence of all 20 canonical amino acids in vitro. EICs from LC-MS analyses are shown for each di- or tripeptide. Reactions contained all 20 canonical amino acids either with or without MboC. (*A*) Dipeptides formed in the presence of MboC (red spectra). Black spectra are EICs from the no

MboC control reaction. (*B*) Tripeptides formed from MboC and the Arg-Ala dipeptide (red spectra). Black spectra are EICs from the no MboC control reaction. (*C*) Tripeptides formed from MboC and the hArg-Ala dipeptide (red spectra). Black spectra are EICs from the no MboC control reaction.

To obtain general rules for ATP-grasp enzyme substrate specificity, sequence similarity networks (SSN) were generated. Amino acid sequences were obtained from NCBI and analyzed using the EFI-EST web-tool. When the ATP-grasp enzymes from the phevamine, mangotoxin, tabtoxin, dapdiamide, rhizocticin, and bacilysin biosynthetic pathways were the only enzymes included, the grouping appeared to correlate with the producing strain, suggesting phylogeny is the dominant factor for ATP-grasp sequence (Fig. 2.9A-C). Significantly increasing the E-value for network generation resulted in slightly more edges between nodes (Fig. 2.9A). A larger network was then generated using a BLAST search of all ATP-grasp enzymes. To visualize this network on a regular computer, stringent parameters needed to be set using the program Cytoscape. Interestingly, phylogeny still appeared to be the most dominant factor in ATP-grasp sequence. Also, the substrate scope within a single ATP-grasp enzyme can vary drastically, for example the ATP-grasp enzyme TabS involved in tabtoxin biosynthesis can generate 136 different dipeptides.[10] Therefore, many more ATP-grasp enzymes need to be studied and characterized in terms of their structures and substrate scope so that an ATP-grasp SSN can be used for predicting substrates.

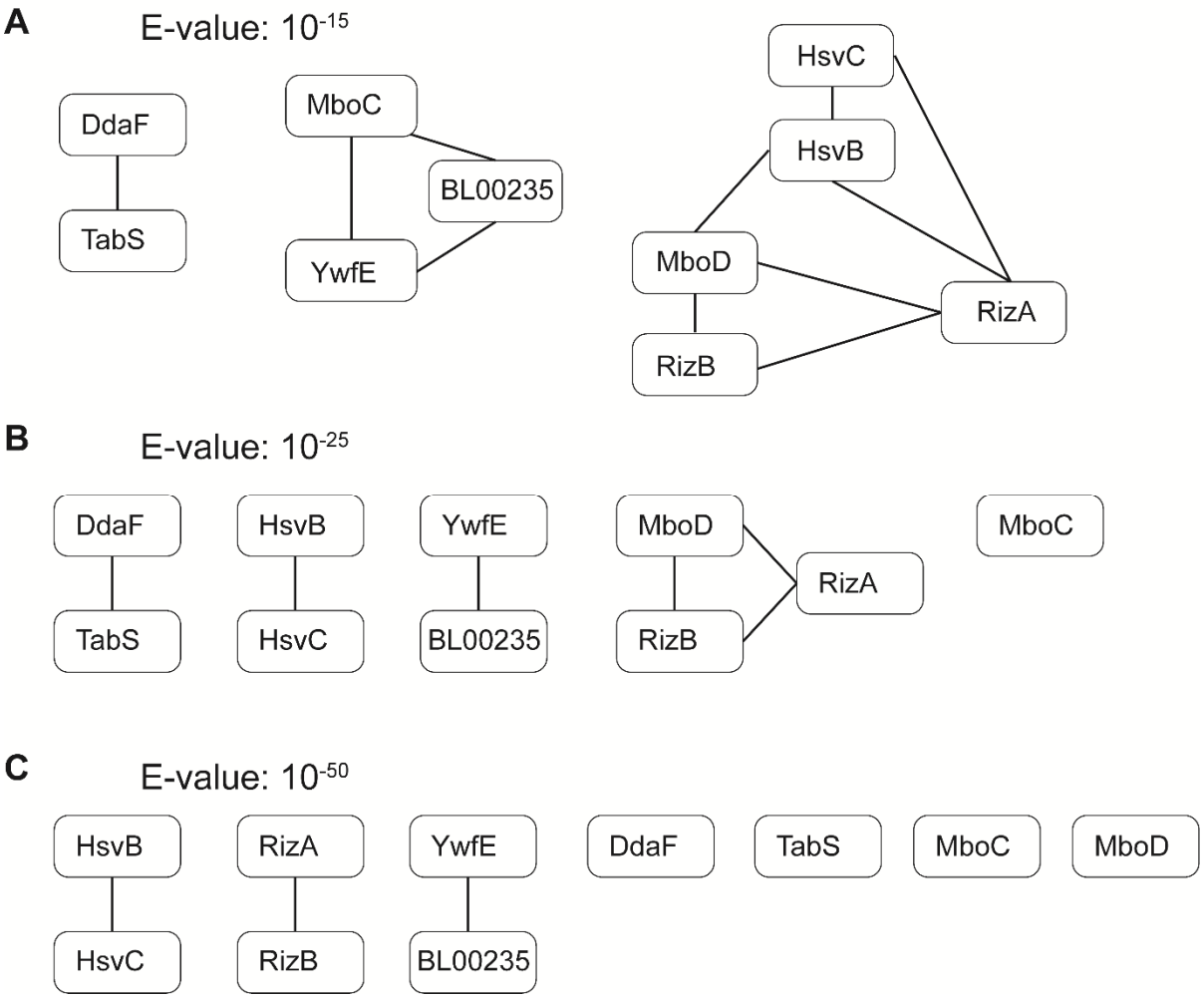


Figure 2.9. ATP-grasp sequence similarity networks group by phylogeny. Sequence similarity networks were generated using EFI-EST and were visualized in Cytoscape. When the E-value was decreased, the number of edges between nodes also decreased. HsvB and HsvC are in the phevamine pathway, MboC and MboD are in the mangotoxin pathway, RizA, RizB, and BL00235 are in the rhizocticin pathway, YwfE is in the bacilysin pathway, DdaF is in the dapdiamide pathway, and TabS is in the tabtoxin pathway.

2.2.8 Activity-guided fractionation of UMAF0158 extracts suggests mangotoxin is unstable.

Mangotoxin is known to inhibit plant ornithine N-acetyl transferase (OAT), thus this bioactivity can be used to track the molecule for extraction and purification. One way to track OAT activity is by *E. coli* growth inhibition, because *E. coli* harbors an *N*-acetylornithinase

that is inhibited by mangotoxin, and this effect can be reversed by supplementation with ornithine.[15] A second way to assess OAT activity is through an in vitro assay using crude plant extract containing the OAT enzyme. The reaction catalyzed by OAT is reversible and involves transfer of an acetyl group from *N*-acetyl-L-Glu (NAcGlu) to L-Orn or *N*-acetyl-L-Orn (NAcOrn) to L-Glu. Therefore, incubating plant OAT with L-Glu and NAcOrn will form L-Orn and NAcGlu as products. Addition of the UMAF extract containing mangotoxin to the reaction will inhibit the activity of OAT and prevent formation of Orn and NAcGlu. Based on this activity assay design, we grew UMAF cultures and extracted polar metabolites, since mangotoxin is likely a short peptide, and fractionated the extract by preparative HPLC. Each fraction was analyzed using both OAT activity detection methods. Indeed, we found that the UMAF overnight seed cultures inhibited *E. coli* growth and plant OAT activity (Fig. 2.10). The culture of *Pto*, which does not produce mangotoxin, was used as a negative control and did not exhibit inhibitory activity against *E. coli* or plant OAT (Fig. 2.10). However, when the UMAF cell-free supernatant, extract, or any of the purified fractions was added, no inhibition of *E. coli* growth or plant OAT was observed, suggesting these materials no longer contain active mangotoxin. Purified mangotoxin was reported to be unstable and inactive after a few hours at room temperature, which likely explains why only the UMAF whole cell culture could inhibit *E. coli* growth and plant OAT activity and purified fractions could not.[15] Perhaps the oxidoreductase MboB is responsible for an oxidation that causes mangotoxin to be unstable under laboratory conditions, which may not be a perfect mimic of the conditions where the *mbo* host, *P. syringae*, is isolated from. To isolate and purify mangotoxin for structural characterization, we will need to identify conditions for culture growth, metabolite extraction, and purification to minimize mangotoxin degradation.

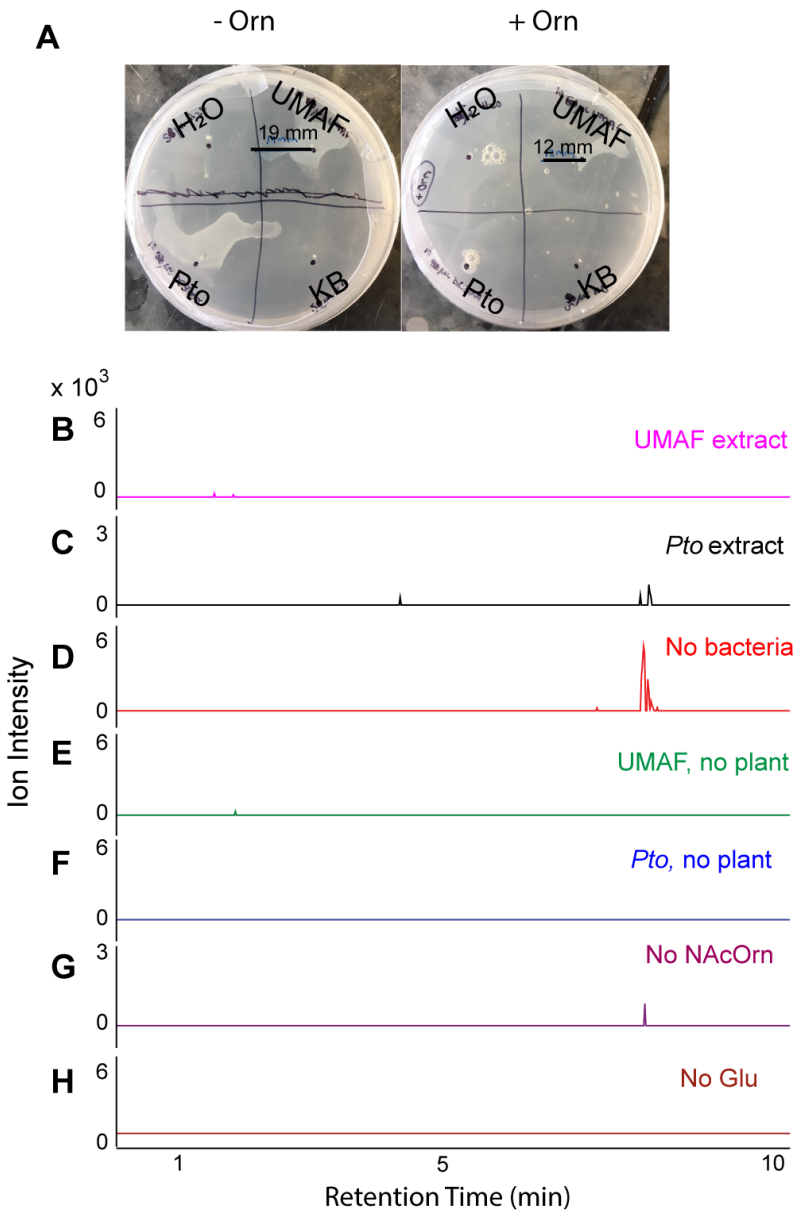


Figure 2.10. Mangotoxin inhibits E. coli growth which can be partly rescued by ornithine and plant OAT activity. (A) Growth inhibition assays showing UMAF extract inhibits *E. coli* growth, but *Pto* extract cannot. No growth inhibition was observed in water and KB media controls. Zones of inhibition radii are measured and shown for UMAF extracts. When top agar was supplemented with Orn (right), the zone of inhibition was reduced in size. (*B*-*H*) EICs of the *N*-acetyl-L-Glu product are shown to assess plant OAT activity under various assay conditions. (*B*) Full assay containing UMAF extract. (*C*) Full assay with DC3000 extract, which does not make mangotoxin. (*D*) Assay without either bacterial extract. (*E*) Assay containing UMAF bacterial extract but no plant extract. (*F*) Assay containing *Pto* bacterial

extract but no plant extract. (*G*) Assay without N-acetyl-L-Orn substrate. (*H*) Assay without L-Glu substrate. The reaction product is only observed in (*C*), (*D*), and (*G*).

2.3 CONCLUSION AND OUTLOOK

We identified a new bacterial virulence factor, phevamine A. It is a small molecule, the biosynthesis of which is controlled by the HrpL virulence regulator in *P. syringae*. Interestingly, phevamine A shares structural similarity with insect polyamine toxins: argiotoxin from the orb-weaver spider, and philanthotoxin from the Egyptian solitary wasp,[86] both of which contain a polyamine and two amino acids, and are powerful neurotoxins that target ion channels. The structural similarity suggests related mechanisms, which will be the focus of future studies.

The importance of phevamine A in bacterial virulence has been demonstrated in two different pathogens, *E. amylovora,*[53] and *P. syringae* in this study. We show that the *hsv* operon and purified phevamine A can suppress both early and late markers of plant immune responses. Specifically, the *hsv* operon suppresses T3SS/harpin-induced callose deposition and host cell death. The phylogenetic distribution of the *hsv* operon suggested that phevamine A is a conserved virulence factor among bacterial phytopathogens. The anti-correlation of *hsv* and the biosynthetic genes for phaseolotoxin within a closely related clade of *P. syrinage* pv. *actinidae* genomes suggested functional redundancy between phevamine A and phaseolotoxin. We demonstrated that phevamine A impairs plant immune signal potentiation by polyamines and arginine, while it is known that phaseolotoxin targets the synthesis of polyamines and arginine.

Our discovery of phevamine A also provided the opportunity to describe new functions for polyamines in the potentiation of the MAMP-induced ROS burst induced by flg22,

expanding the role of polyamines as important defense mediators. Our results are consistent with a model (Fig. S2.28) in which *P. syringae* uses either phaseolotoxin or phevamine A to dampen host immune output by altering polyamine biosynthesis or the signaling capacity of polyamines, respectively. We speculate that components of the T3SS and/or harpins may be recognized at the cell surface, acting as elicitors of MTI subject to polyamine potentiation and suppression by phevamine A.

Further work will be conducted to characterize the structure of mangotoxin and its biosynthesis. The entire *mbo* cluster has been cloned into an *E. coli* overexpression vector and metabolomics analyses were conducted comparing overexpression with empty vector yielding several masses of interest, but it is unclear which is/are biologically relevant molecules. Activity of the oxidoreductase MboB will be reconstituted in vitro, likely under anaerobic conditions, and other amine acceptors, including the canonical amino acids, will be screened for MboE amidinotransferase activity. Concurrently, molecules of interest from metabolomics experiments will be analyzed and compared with enzyme assay products to identify the mass of mangotoxin. Once masses of interest are identified, several rounds of purification from in vitro enzyme assays or culture extract will be conducted in a single day to isolate rapidly pure compound for structural analysis by MS/MS and NMR. The purified product will then be examined in biological assays, including OAT inhibition and *E. coli* growth inhibition, but again should be conducted immediately after purification to avoid molecule degradation.

We show that two ATP-grasp enzymes encoded in the *hsv* cluster are responsible for assembling phevamine A. Preliminary results suggest that two different ATP-grasp enzymes encoded in the *mbo* cluster are responsible for mangotoxin production. These enzymes belong to a large superfamily of enzymes that catalyze amide bond formation by activating carboxylic

acids through phosphorylation.[87] ATP-grasp-type enzymes are widely used by bacteria for the synthesis of diverse small molecules, including the virulence factors phaseolotoxin and dapdiamide.[11, 16] Studies of these enzymes in a marine proteobacterium and a soil actinomycete have led to the discovery of small molecules with distinct structures but unknown functions.[38, 88] Still, most biosynthetic operons containing ATP-grasp-type enzymes have eluded genome-guided discovery. Our study suggests that targeted mining of uncharacterized ATP-grasp enzymes will likely lead to novel small molecules.

Our approach for identifying cryptic small molecules builds on the rich knowledge of pathogen gene expression and integrates comparative genomic analysis, biochemical enzyme reconstitution, and physiological assays. Bacterial genomes harbor many biosynthetic operons of unknown function; therefore, this approach holds potential for identifying many more new small molecules essential for bacteria–host interactions.

2.4 EXPERIMENTAL PROCEDURES

2.4.1 General

All chemicals were purchased from Sigma-Aldrich or Fisher. Primers were ordered from Integrated DNA Technology (Table S2.2). Polymerase, restriction enzymes, and ligase were purchased from New England Biolabs. DNA manipulation kits were purchased from Qiagen and Thermo Scientific. PCR reactions were carried out using an Eppendorf Nexus GSX1 thermocycler with Q5 DNA polymerase. General genetic manipulation of *E. coli* was carried out using standard protocols.[89] LC-MS data was obtained using an Agilent 6520 accurate-mass Q-TOF LC/MS. Preparative HPLC was carried out using a Phenomenex Luna 10 μm C18 (2) column on a Varian Prostar instrument equipped with a PDA detector. Further

purification was carried out using a Phenomenex Kinetex 5 µm C18 column (100 Å, 150 mm x 4.60 mm) on a Shimadzu HPLC equipped with an SPD-20A UV-vis detector. NMR spectra were obtained using a Bruker 600 MHz spectrometer or a Bruker Avance^{III} HD 800 MHz spectrometer unless otherwise noted.

2.4.2 Bacterial strains and culture conditions

For cloning, *E. coli* DH5α and *E. coli* Top10 chemically competent maximum efficiency cells were used (Invitrogen). For heterologous expression, *E. coli* BL21(DE3) chemically competent cells (Agilent) were transformed with vectors containing *hsv* or *mbo* expression constructs (Table S2.3). The *hrpL* expression construct in the pBAD vector was previously described.[13] All overnight seed cultures were inoculated with a single colony and grown in 5 mL media in 14 mL Falcon culture tubes. Antibiotics were used in these concentrations, unless otherwise noted: 10 µg/mL tetracycline, 25 µg/mL chloramphenicol, 50 µg/mL kanamycin, 50 µg/mL spectinomycin, or 100 µg/mL ampicillin. All cultures for metabolomics were grown in 100 mL media with the appropriate antibiotic and 100 µL−2 mL of a saturated overnight culture. All *E. coli* cultures were grown in LB Media (Miller, granulated, Fisher) at 37 °C to an OD_{600} between 0.4 and 0.8. *E. coli* cultures were induced with 0.5 mM IPTG and then were incubated at 16 °C overnight, shaking at 225 rpm. *P. syringae* and *P. fluorescens* starter cultures were grown overnight in King's B Media (20 g/L Bacto Proteose Peptone #3, 1% glycerol by volume, 1.5 g/L K_2HPO_4, and 0.734 g/L $MgSO_4$, dissolved in water). *P. syringae* cultures containing the pBAD plasmid were induced with 0.2% L-arabinose. All other *P. syringae* pv. *tomato* DC3000, *P. syringae* pv. *syringae* UMAF0158,[14] and *P. fluorescens* cultures were not induced. Wildtype *P. fluorescens Pf*0-1 harboring the *hsv*-

pBBR5 vector was grown in the presence of 50 µg/mL ampicillin and 50 µg/mL gentamycin. *P. fluorescens* harboring both *hsv*-pBBR5 and the T3SS (*Pf*0-1 +T3SS) was grown with 25 µg/mL tetracycline and 50 µg/mL gentamycin. After reaching an OD_{600} of >0.6 in King's B media, *P. fluorescens* and *P. syringae* cultures were washed with 100 mM $MgCl_2$ and resuspended in minimal media (12.8 g/L Na_2HPO_4 ● 7 H_2O, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L NH_4Cl, 1% glycerol by volume, 0.024 g/L $MgSO_4$, 0.0011 g/L $CaCl_2$, and 0.18 g/L mannitol, dissolved in water). Once resuspended in minimal media, these cultures were grown at 28 °C overnight and 16 °C for another 24 hours shaking at 225 rpm.

2.4.3 Metabolite extraction for metabolomics

E. coli, P. syringae, and *P. fluorescens* cultures (100 mL) were spun down at 3500 x *g* and 4 °C for 10 minutes. Solvent extractions were performed using an equal volume of supernatant to chloroform-methanol (2:1:1). The top aqueous layer was separated into a round-bottom flask and concentrated under vacuum. The dried down material was resuspended and transferred to a small glass vial, concentrated under reduced pressure, and stored at −20 °C until MS analysis.

2.4.4 Protein expression and purification

Large-scale cultures for protein purification were grown in 1 L of media containing the appropriate antibiotic and 1−5 mL of overnight culture. At an OD600 of 0.5−0.8, a sample of 0.5 mM IPTG was added to induce protein expression. After 16 hours of protein expression, cultures were spun down and resuspended in HEPES lysis buffer (50 mM HEPES, 150 mM NaCl, 10% glycerol, pH 7.5). Cells were lysed by sonication at 30% amplitude for 3 minutes

(0.5 seconds on and 1.5 seconds off). Cell debris was removed by centrifugation at 17,000 rpm for 40 minutes and the supernatant was filtered using a 0.45 µm filter (Corning). Each Hsv protein was purified using an AKTA FPLC using a 5 mL HiTrap nickel column (GE) at a flow rate of 3.0 mL/min. Wash buffer consisted of 50 mM HEPES, 150 mM NaCl, 25 mM imidazole, and 10% glycerol (pH 7.5). Elution buffer was the wash buffer with the addition of 250 mM imidazole. The wash step lasted for 5 column volumes (CVs). The proteins were eluted using a gradient of 0−100% of elution buffer over 10 CVs, and then held at 100% for 5 CVs. HsvB, HsvC, and MboD were further purified using a HiLoad 16/600 Superdex 200 prep grade size exclusion column (GE) on the FPLC using the lysis buffer described above. The column was equilibrated with lysis buffer at a flow rate of 0.5 mL/min for 1.5 CVs, then protein was loaded onto the column and eluted over 1.5 CVs using the same buffer and flow rate.

2.4.5 Amidinotransferase substrate assay conditions

To identify the preferred substrates of the amidinotransferase proteins HsvA and MboE, L-arginine was incubated with potential substrates that accept the amidino group. Each reaction mixture contained two amino acids at 10 mM each (L-Arg and L-Lys, L-Arg and spermidine, or L-Arg and spermine). For MboE, L-Cys, Gly, and L-Ala were also tested with L-Arg. Each mixture also contained 100 mM HEPES (pH 7.5) and 20 µM purified HsvA to a final volume of 50 µL. Samples were incubated at 37 °C for 2 hours. Enzymes were omitted in negative controls. A sample of 10 µL was injected onto a Gemini C18 column (Phenomenex) on LC-MS and separated with a gradient of 2−73% acetonitrile (0.1% formic acid) over 25 minutes. This assay showed that HsvA catalyzes the transfer of the amidino group from L-Arg to spermidine, spermine, or L-lysine. Based on peak areas, the amidinospermidine product is

present at the highest level, suggesting spermidine is the preferred substrate. MboE activity could not be reconstituted.

2.4.6 ATP-grasp substrate assay conditions

A malachite green dye assay (P_iColorLock, Innova Biosciences) was used to identify the preferred substrates of the ATP-grasp enzymes HsvB and MboC by quantifying free phosphate (P_i) release. Activity of MboD and HsvC were not reconstituted using this method. Substrate assays were conducted in 100 µL scale in a microcentrifuge tube, and consisted of 2.0 mM amino acids, 10 µM of each enzyme, 10 mM $MgCl_2$, 5 mM ATP, and 50 mM HEPES (pH 7.5). After 30 minutes, reactions were diluted 1 in 10 with DI water into a 96-well plate to a final volume of 200 µL, then 50 µL Gold mix reagent and 20 µL stabilizer were added. After 30 minutes of incubation, UV-vis absorption of the products at 635 nm was measured on a Tecan infinite M1000 Pro plate reader. Assays containing L-Phe produced the most P_i, indicating that HsvB preferentially activates L-Phe. MboC activity was observed when hArg and Ala, hArg and Cys, and Arg and Ala were incubated together.

2.4.7 Characterization of phevamine biosynthetic pathway

Sequential assays were conducted to elucidate the biosynthesis of phevamine A. Based on the structures of pre-phevamine, amidinotransfer by HsvA or condensation with L-Val by HsvC may occur as the first step. This rationale led us to examine three potential biosynthetic pathways: HsvA → C → B, HsvC → A → B, and HsvC → B → A. An initial assay was conducted for either HsvA or HsvC. The HsvA assay consisted of 10 mM L-arginine, 10 mM spermidine, 100 mM HEPES (pH 7.5), 20 µM HsvA, and water in a 100 µL volume. Protein

was removed after a 30 minute incubation at room temperature using a 3000 Da Amicon Ultra filter (Millipore) and the small molecule filtrate was collected. To this assay mixture, we added 2 mM $MgCl_2$, 1 mM ATP, 2 mM L-valine, and 10 μM HsvC. Protein was again removed after 30 minutes at room temperature with a 3000 Da filter and an aliquot was collected for LC-MS analysis. Finally, 2 mM L-phenylalanine and 10 μM HsvB were added, and the reaction was incubated at room temperature for 30 minutes before the reaction was quenched with acetonitrile at a ratio of 1:1 for 15 minutes, precipitated protein was removed by centrifugation, and the supernatant was analyzed by LC-MS. The initial HsvC assays consisted of 2 mM spermidine, 2 mM L-valine, 100 mM HEPES (pH 7.5), 2 mM $MgCl_2$, 1 mM ATP, 10 μM HsvC, and water in a 100 μL volume. After 30 minutes at room temperature, HsvC was removed with a 3000 Da filter. To one set of assays, 10 mM L-arginine and 20 μM HsvA were added. To the other set of assays, 2 mM L-phenylalanine and 10 μM HsvB were added. Both sets of reactions were incubated at room temperature for 30 minutes before removing protein with a 3000 Da filter. Aliquots were taken for LC-MS analysis. These reactions were then incubated with the remaining constituents that had not been added: 2 mM L-phenylalanine and 10 μM HsvB or 10 mM L-arginine and 20 μM HsvA. Reactions were incubated at room temperature for 30 minutes before they were quenched with acetonitrile at a ratio of 1:1 and the precipitated protein was removed by centrifugation. All aliquots were diluted 5 folds with water for LC-MS analysis. Phevamines were only produced when the reaction order was HsvA → HsvC → HsvB (Figure S2.10).

2.4.8 Preparative scale in vitro enzymatic synthesis of phevamines

One milliliter one-pot enzymatic assays for phevamine A were conducted at room temperature for two hours. Assays contained 5 mM L-arginine, 5 mM spermidine, 1 mM L-valine, 0.5 mM L-phenylalanine, 100 mM HEPES (pH 7.5), 2 mM $MgCl_2$, 1 mM ATP, 20 μM HsvA, 10 μM HsvB, 10 μM HsvC, and water. The reaction was then quenched with 1 mL acetonitrile to precipitate protein overnight at −20 °C. The precipitated proteins were removed by centrifugation. Phevamines were purified from the supernatant using one round of preparative HPLC, and two rounds of analytical HPLC using conditions described in the supplementary methods.

2.4.9 Purification of phevamine A

Assay supernatants were concentrated under reduced pressure and injected onto the prep HPLC at a flow rate of 15 mL/min with water and acetonitrile (both containing 0.1% trifluoroacetic acid) as mobile phases. The solvent gradient was held at 2% acetonitrile for 10 minutes, then 5% acetonitrile for 5 minutes, and 10% acetonitrile for 15 minutes, before being ramped up to 95% over another 20 minutes. Fractions containing phevamine A (28–30 min retention time) were concentrated under reduced pressure and injected onto the analytical HPLC using the same mobile phases as those for the prep HPLC. The solvent gradient ramped from 2−30% acetonitrile over 35 minutes before ramping up to 95% for another 8 minutes (flow rate: 0.5 mL/min). Fractions containing phevamine A were collected at retention times of 20−25 minutes, and concentrated. To separate the two regio-isomers, a final round of purification was conducted on the analytical HPLC with an extended gradient at 10−20% acetonitrile for 25 min. Specifically, the solvent gradient started at 2−5% acetonitrile over 5

minutes, then 5−10% over another 5 minutes, and finished with 10−20% for another 25 minutes (flow rate: 0.5 mL/min). Fractions with retention times between 17.5–19.5 minutes were collected and concentrated. LC/MS analysis showed that the concentrated sample contained ~95% pure phevamine A (67% yield).

2.4.10 NMR spectroscopy for phevamine A and B

All spectra for in vitro purified phevamine A and B were obtained using a Bruker Avance^{III} HD (800 MHz ^{1}H reference frequency, 201 MHz for ^{13}C) equipped with a 5 mm CPTCL ^{1}H-^{13}C/^{15}N cryo probe. Non-gradient phase-cycled dqfCOSY spectra were acquired using the following parameters: 0.6 s acquisition time, 400−600 complex increments, 8, 16 or 32 scans per increment. Non-gradient HSQC and HMBC spectra were acquired with these parameters: 0.25 s acquisition time, 200−500 complex increments, 8−64 scans per increment. ^{1}H, ^{13}C-HMBC spectra were optimized for $J_{H,C}$ = 6 Hz. HSQC spectra were acquired without decoupling. NMR spectra were processed and baseline corrected using Mestrelabs MNOVA software packages. NMR assignments are shown in Tables S2.4 and S2.5.

2.4.11 MS and MS/MS analysis

Culture extracts, in vitro synthesized phevamines, and chemically synthesized phevamine A, were analyzed by Agilent 6520 accurate-mass Q-TOF LC/MS using a Phenomenex Kinetex 5 µm C18 column (100 Å, 150 mm x 4.60 mm). Mobile phases were water and acetonitrile, each containing 0.1% formic acid. The gradient was held at 2% acetonitrile for 2 minutes before ramping up to 45% acetonitrile over 17 minutes at a flow rate of 0.4 mL/min. For each phevamine sample, *m/z* 434.324 was selected at a retention time of 8

± 5 minutes and analyzed by MS fragmentation. Targeted MS/MS was conducted using 30 V collision energy, 970.9 ms/sec acquisition time, and a 4 *m/z* isolation width.

2.4.12 Plant material and bacterial growth

Arabidopsis thaliana was grown in walk-in growth rooms maintained at 21 °C/18 °C (day/night) with a 9 hour/15 hour (day/night) cycle. *Nicotiana benthamiana* was grown in a walk-in growth room maintained at 26 °C/22 °C with a 12 hour/12 hour (day/night) cycle and a LGM550 professional LED grow light system (LED Grow Master Global). For maintenance and transformation, *Pseudomonas* strains were grown in King's B media (KB) at 28 °C. For infiltration or dipping *in planta*, *Pseudomonas* strains were grown in liquid culture overnight with the appropriate antibiotics, then plated on a petri dish and incubated overnight before resuspension in 10 mM $MgCl_2$.

2.4.13 Bacterial growth assay

Pto, *Pto*Δ*hsv*, *Pto*-Cor$^-$, and *Pto*-Cor$^-$Δ*hsv* cells were resuspended in 10 mM $MgCl_2$ to a final concentration of 2×10^5 cfu/mL. Twenty-day-old *A. thaliana* Col-0 seedlings were dipped in bacterial solutions supplemented with Silwet L-77 (Momentive), and growth was assessed at day 0 and 3 days post-infection as described by Tornero and Dangl.[90] Eight samples containing 3 seedlings were collected for each treatment.

2.4.14 Callose deposition measurement

Three- to five-week-old *Arabidopsis* plants were infiltrated with *Pseudomonas* strains at an OD_{600} of 0.2 in 10 mM $MgCl_2$ and collected after about 20 hours. To visualize callose

deposition, leaves were stained with aniline blue.[91] The tissue was cleared and dehydrated with 96% ethanol overnight at 37 °C. Cleared leaves were washed with distilled water and then stained in 0.01% aniline blue in 150 mM K_2HPO_4 (pH 9.5) for 4 hours at room temperature. Stained samples were washed and mounted in distilled water and examined by epifluorescence (LEICA M205 FA) with 100X magnification. Images were taken at the region below the infiltrated zone of each leaf. Counting of accumulated callose foci was carried out using ImageJ (NIH). For each treatment, 10 to 20 leaves were processed.

2.4.15 Phylogenetic Analysis

The phylogenetic history of *Pseudomonas syringae* was inferred by Yang *et al.*[73] Protein sequences of toxin biosynthetic genes listed in Table S2.1 were used as BLAST queries to search for homologous protein sequences in the 292 *P. syringae* genomes (available as of Jan. 2017, PATRIC). A hit with over 80% protein sequence identity was considered positive. The tree was visualized using iTOL.[92]

2.4.16 Metabolite dependence on T3SS

The *hsv* operon was cloned into the pBBR5 plasmid containing a constitutive promoter and gentamycin resistance marker.[93] *P. fluorescens Pf*0-1 and *Pf*0-1 + T3SS were both transformed with *hsv*-pBBR5 by electroporation using the method described by Choi and Schweizer.[94] Cultures (100 mL) were grown and extracted as described above to detect phevamine production and determine whether phevamine secretion is dependent on T3SS (Fig. S2.7).

2.4.17 Generation of Pseudomonas clean deletion mutants

Pseudomonas knockout clones were generated using MTN1907, a modified version of pLVC-D, which allows for *SacB* counter-selection.[13, 95-96] Overlapping primers MT1881, MT1880, MT1877, and MT1878 (Table S2.2) were used to generate a chimeric fragment composed of the ~200 bp sequence upstream of *hsv*A fused to the ~200bp region downstream of *hsv*C. The chimeric fragment was sub-cloned into MTN1907 and *Pseudomonas* was transformed with the plasmid to obtain a merodiploid after recombination. Two merodiploids, with a 3' end and a 5' end insertion, were selected to generate two independent *Pseudomonas* clones with clean *hsv* deletion for each parental strain (*Pto* DC3000, *Pto*-Cor $^{-}$, *Pto*Δ*hrcC*, *Pto*D28E) after growth on media containing 5% sucrose for the second recombination (see Fig. S2.29 and S2.30). For complementation, the *hsv* operon was sub-cloned into pBAV226 downstream of the *NptII* promoter and transformed into *Pto*D28EΔ*hsv2* (see Fig. S2.30). The pBAV226 vector contains the sequence encoding the HA-tag downstream of the gateway cassette, but the HA tag was not included in the product of *hsv*C because the stop codon was retained.

2.4.18 ROS burst measurement

Leaf disks from 4-week-old *A. thaliana* Col-0, or *N. benthamiana* were placed into a 96-well plate with 100 µl of water in each well. Twelve leaf disks were used per treatment. After overnight incubation for *Arabidopsis* leaf disks, or 20 hours for *N. benthamiana* leaf disks, each sample was treated with 100 µl of reaction mix, including 17 mg/mL of luminol (Sigma), 10 mg/mL of horseradish peroxidase (HRP; Sigma), distilled water, flg22 at concentrations ranging from 5 nM to 100 nM, and L-arginine, and polyamines at concentrations

typically ranging from 200 μM to 800 μM. Phevamines and pre-phevamine were used at the same concentration as L-arginine (Sigma) or spermidine. Luminescence was measured immediately with 0.5 s integration and 1 min interval over 45 min using a SpectraMax L (Molecular Devices). For each experiment, 8 to 12 leaf-disks per condition were monitored. The flg22 peptide (QRLSTGSRINSAKDDAAGLQIA) was synthesized by Genscript.[78]

2.4.19 Calcium burst measurement

Leaf disks of 3- to 4-week-old *A. thaliana* lines pMAQ2 expressing the apoaequorin gene under the control of the cauliflower mosaic virus 35S promoter[85] were collected in a 96-well plate and incubated overnight with 100 μM coelenterazine native (BYOSYNTH).[85] Luminescence was measured as in the ROS burst assay, with the exception that luminol and HRP were excluded from the reaction mix.

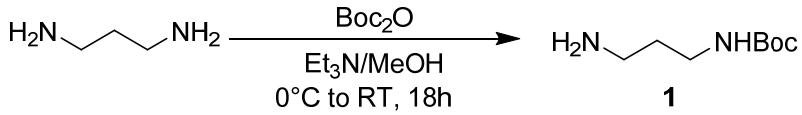
2.4.20 Synthesis of phevamine A

General information

Chemicals were purchased from commercial suppliers and used without further purification. Fmoc-derivatized amino acids were purchased from Alfa Aesar. N^1, N^4-bis-boc-spermidine was purchased from Chem-Impex International. Diethylamine and DIPEA were purchased from Sigma-Aldrich. *N,N′*-Di-Boc-1H-pyrazole-1-carboxamidine and trifluoroacetic acid were purchased from Acros Organics. HATU was purchased from Oakwood Products Inc. NMR solvents were purchased from Cambridge Isotope Laboratories. Normal-phase column chromatography was performed on silica gel (230−400 mesh). Reversed-phase chromatography was performed on C18-bonded silica, either on Phenomenex

Kinetex C18 or Phenomenex Luna C18. High-resolution mass spectra (HRMS) were obtained using an Agilent 6520 Q-TOF LC/MS. Proton, carbon and fluorine magnetic resonance spectra (^{1}H NMR, ^{13}C NMR, and ^{19}F NMR) were recorded on a Bruker model DRX 400 or 600 (^{1}H NMR at 400 MHz or 600 MHz, ^{13}C NMR at 101 MHz or 151 MHz) or on a Bruker AVANCE III-OneBay500 (^{13}C NMR at 235 MHz) spectrometer. NMR experiments are reported in δ units, parts per million (ppm), and were referenced to $CDCl_3$ (δ 7.26 ppm for ^{1}H and 77.0 ppm for ^{13}C), DMSO (2.50 ppm), MeOD (3.31 ppm) or D_2O (4.79 ppm) as internal standards. ^{1}H NMR data are reported as follows: chemical shift, multiplicity (s = singlet, bs = broad singlet, d = doublet, dd = doublet of doublets, t = triplet, td = triplet of doublets, m = multiplet), coupling constants (Hz), and integration.

_N_-Boc-1,3-diaminopropane (1)



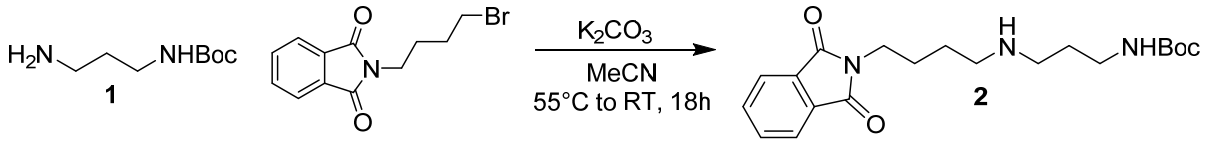
1

This synthesis followed a similar procedure as Kaur *et al.*[97] 1,3-Diaminopropane (3 g, 40.5 mmol, 1 eq.) was dissolved in a 10% solution of Et_3N in MeOH (30 mL); the mixture was cooled to 0 °C. In a flame dried flask, di-*tert*-butyl dicarbonate (4.4 g, 20 mmol, 0.5 eq.) was dissolved in MeOH and added to the cooled solution (via addition funnel). The reaction mixture, which did not appear homogeneous, was stirred overnight at room temperature (RT). The solvent was then removed under reduced pressure. The crude mixture was treated with 10% Na_2CO_3 (10 mL, three times) and extracted with CH_2Cl_2. The combined organic phases were dried over anhydrous Na_2SO_4 and the solvent was removed under reduced pressure. The

desired product (**1**) was obtained as a white solid (1.65 g, 23% yield) and used in the subsequent step without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.20 (bs, 1H), 1.44 (s, 9H), 1.64-1.57 (m, 2H), 2.76 (t, J = 6.6 Hz, 2H), 3.22-3.16 (m, 2H), 4.90 (bs, 1H).

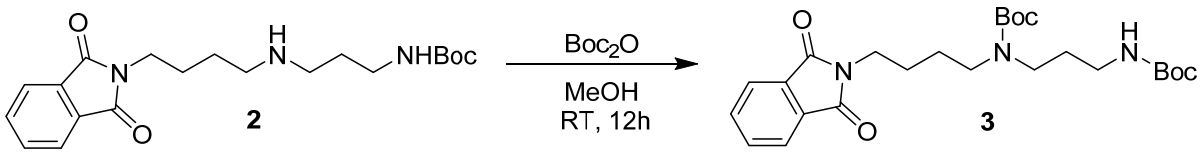
***N*-(4-(3-Boc-1,3-diaminopropane)butyl)phthalimide (2)**



This reaction followed a procedure similar to that used by Wang *et al.* (2012).[98] *N*-Boc-1,3-diaminopropane (1.38 g, 7.9 mmol, 1.1 eq.) was dissolved in MeCN (15 mL) at RT and K_2CO_3 (3.43 g, 24.8 mmol, 5 eq.) was added. *N*-(4-Bromobutyl)phthalimide (2.2 g, 7.7 mmol, 1 eq.) was then added and the reaction mixture was heated to 50 °C and stirred at this temperature for 18 hours. Subsequently, the mixture was warmed up to RT and treated with H_2O. The mixture was extracted with EtOAc; the combined organic phases were dried over anhydrous Na_2SO_4 and the solvent was removed under reduced pressure. The desired product (**2**) was obtained (3 g, quantitative yield) and used in the subsequent step without further purification.

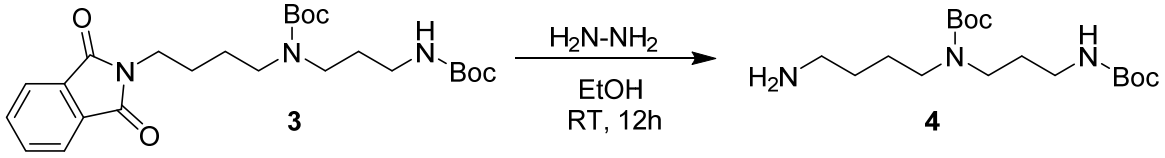
^{1}H NMR (400 MHz, $CDCl_3$) δ 1.90−1.82 (m, 2H), 1.40 (bs, 9H), 1.65-1.56 (m, 2H), 1.75−1.69 (m, 2H), 2.39−2.36 (m, 2H), 3.20-3.13 (m, 4H), 3.71−3.64 (m, 2H), 4.96 (bs, 1H), 7.75−7.72 (m, 2H), 7.82−7.79 (m, 2H),

***N*-(4-(*N*^{1},*N*^{3}-bis-Boc-1,3-diaminopropane)butyl)phthalimide (3)**



This reaction followed a similar procedure to that was used by Wang *et al.* (2014).[99] The substrate was dissolved in methanol (40 mL) and di-*tert*-butyl dicarbonate (1.7 g, 7.8 mmol, 1 eq.) was added. The reaction mixture was stirred at RT for 12 hours, then the solvent was removed under reduced pressure. The Boc-bisprotected product (**3**) was obtained as a pale, yellow oil (3.3 g, 87% yield) and used in the subsequent step without further purification. ^{1}H NMR (400 MHz, $CDCl_3$) δ 1.45−1.44 (m, 2H), 1.70−1.67 (m, 5H), 1.95−1.85 (m, 1H), 2.42 (bs, 2H), 3.27−3.10 (m, 4H), 3.76−3.70 (m, 2H), 7.75−7.72 (m, 2H), 7.88−7.84 (m, 2H).

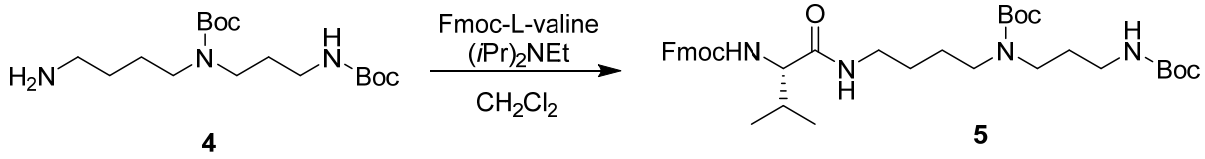
***N*^{1}, *N*^{4}-bis-Boc-spermidine (4)**



This reaction used a procedure similar to that was used by Wang *et al.* (2014).[99] To a solution of *tert*-butyl (3-((tert-butoxycarbonyl)amino)propyl)(4-(1,3-dioxoisoindolin-2-yl)butyl)carbamate (**3**) (780 mg, 1.64 mmol) in EtOH (30 mL), hydrazine hydrate (0.83 mL) was added and the reaction mixture was stirred for 12 hours at RT. The solvent was removed under reduced pressure and the residue was suspended in 10% ammonia and extracted with $CHCl_3$; the combined organic phases were dried on anhydrous Na_2SO_4 and the solvent was removed under reduced pressure. The desired product (**4**) was obtained as a pale, yellow oil (340 mg, 60% yield) and used in the subsequent step without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.48−1.46 (m. 18H), 1.66−1.57 (m, 7H), 2.43−2.40 (m, 1H), 2.49−2.46 (m, 1H), 2.74−2.71 (m, 1H), 3.26−3.12 (m, 4H), 3.69 (bs, 1H), 4.90 (bs, 1H), 5.78, 5.64 (bs 1H).

Fmoc-(*S*)-Val-bis-Boc-spermidine (5)

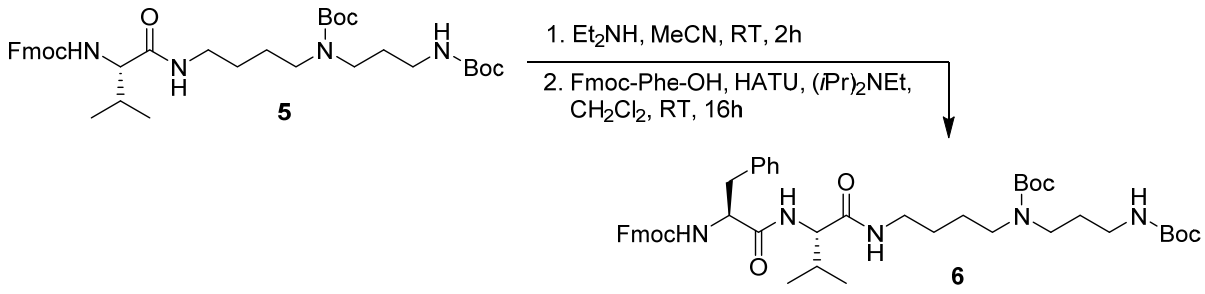


4 5

To a solution of N^1, N^4-bis-Boc-spermidine (**4**) (250 mg, 0.724 mmol, 1.0 eq.) in CH_2Cl_2 (20 mL) on ice, Fmoc-Val-OH (295 mg, 8.68 mmol, 1.2 eq.), and DIPEA (252.2 µL, 1.448 mmol, 2.0 eq.) were added. The reaction was stirred at room temperature for 16 hours. The crude reaction mixture was purified by flash chromatography (hexanes/ethyl acetate gradient) to afford pure product (**5**) (310 mg, 64% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 0.97 (m, 6H), 1.46 (s, 9H), 1.47 (s, 9H), 1.52 (m, 2H), 1.66 (m, 2H), 2.15 (m, 2H), 3.07-3.35 (m, 9H), 4.23 (t, J = 7.0 Hz, 1H), 4.36 (m, 1H) 4.44 (m, 1H), 7.32 (t, J = 7.4 Hz, 2H), 7.42 (t, J = 7.6 Hz, 2H), 7.61 (d, J= 7.6 Hz, 2H), 7.78 (d, J = 7.2 Hz, 2H).

Fmoc-(*S*)-Phe-(*S*)-Val-bis-Boc-spermidine (6)

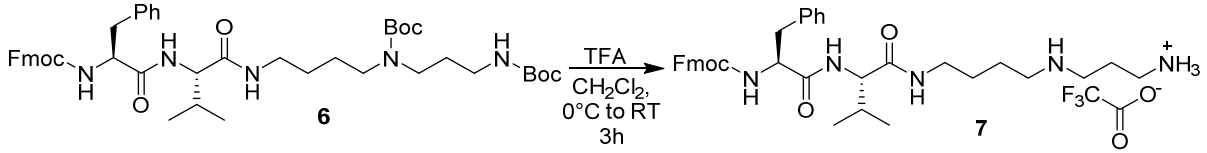


5 6

To a solution of Fmoc-(S)-Val-bis-Boc-spermidine (**5**) (310 mg, 0.465 mmol, 1.0 eq) in 20 mL MeCN, diethylamine (500 □L, 2.5% v/v) was added. The reaction was stirred at room temperature for 2 hours. The crude reaction mixture was concentrated under reduced pressure. Ethyl acetate (25 mL) was added, and the reaction mixture was concentrated again. The crude reaction mixture was dissolved in CH_2Cl_2 (20 mL) and added to a mixture of Fmoc-Phe-OH (216 mg, 0.558 mmol, 1.2 eq.), HATU (269 mg, 0.6975 mmol, 1.5 eq.), and DIPEA (162 µL, 0.742 mmol, 2.0 eq.) previously kept under stirring for 20 minutes on ice. The reaction was stirred at room temperature for 16 hours. The crude reaction mixture was purified by flash chromatography (hexanes/ethyl acetate gradient) to afford pure product (**6**) (190 mg, 50% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 0.87 (d, J = 6.8 Hz, 3H), 0.92 (d, J = 6.8 Hz, 3H), 1.47 (s, 18H), 1.51 (m, 2H), 1.69 (m, 2H), 2.19 (m, 2H), 3.05−3.35 (m, 11H), 4.19 (t, 1H), 4.31 (m, 1H), 4.41 (m, 1H), 4.51 (m, 1H), 7.20 (m, 2H), 7.31 (m, 5H), 7.41 (t, J = 7.4 Hz, 2H), 7.54 (d, J = 7.6 Hz, 2H), 7.77 (d, J = 7.2 Hz, 2H).

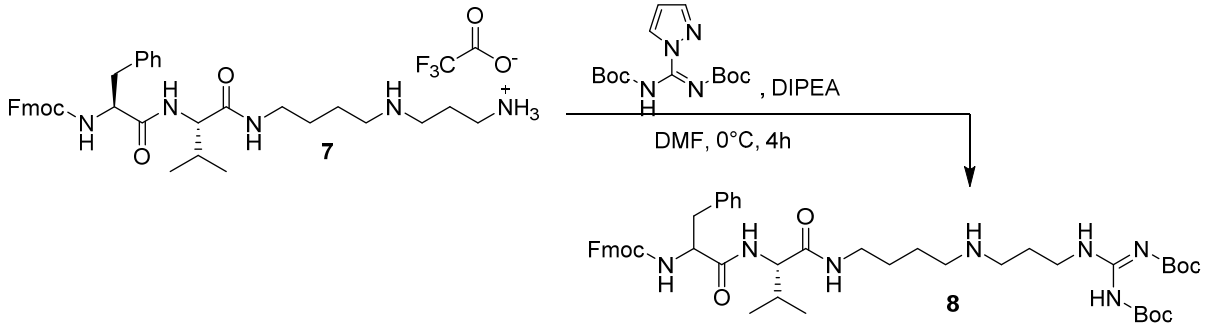
Fmoc-(*S*)-Phe-(*S*)-Val-spermidine (7)



A solution of Fmoc-(*S*)-Phe-(*S*)-Val-bis-Boc-spermidine (**6**) (60 mg, 0.072 mmol) in CH_2Cl_2 (2 mL) was cooled to 0 °C and 800 µL of trifluoroacetic acid was added dropwise. The reaction was stirred at room temperature for 3 hours. The crude reaction mixture was concentrated under reduced pressure. Toluene was added, and the sample was concentrated under reduced pressure until dry. The pure product (**7**) (23 mg, 32% yield over two steps) was

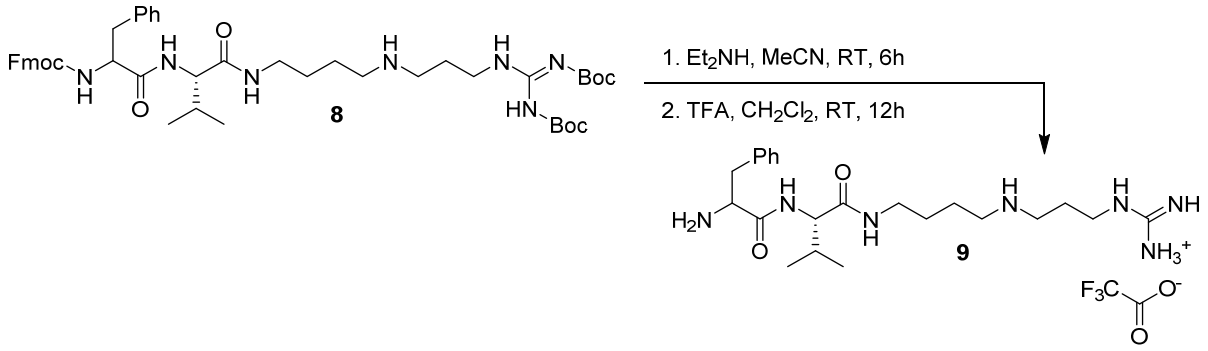
obtained as a white solid: ^{1}H NMR (400 MHz, MeOD) δ 0.96 (m, 6H), 1.59 (m, 2H), 1.70 (m, 2H), 2.07 (m, 3H), 3.00−3.27 (m, 10H), 4.08 (d, J = 7.2 Hz, 1H), 4.16 (t, J = 7.0 Hz, 1H), 4.26 (m, 1H), 4.34 (m, 1H), 4.42 (m, 1H), 7.16 (m, 1H), 7.21−7.32 (m, 5H), 7.40 (t, J = 7.6 Hz, 2H), 7.58 (d, J = 7.6 Hz, 2H), 7.81 (d, J = 7.6 Hz, 2H); HRMS (EI) *m/z* $[M + H]^+$ Calculated for $C_{36}H_{48}N_5O_4$ 614.370, observed 614.371.

Fmoc-(*S*)-Phe-(*S*)-Val-spermidine-*N,N'*-di-Boc-1*H*-guanidine (8)



To a solution of Fmoc-(S)-Phe-(S)-Val-spermidine (**7**) in DMF (2 mL), which has been cooled to 0 °C, DIPEA (200 µL) and *N,N'*-Di-Boc-1H-pyrazole-1-carboxamidine (24 mg, 0.072 mmol, 1.0 eq.) were added. The reaction was stirred for 4 hours at room temperature. The crude reaction mixture was purified directly using reverse-phase prep HPLC (Phenomenex Luna C18, H_2O/MeCN gradient) and then purified using a second round of reverse-phase HPLC (Phenomenex Kinetex C18, H_2O/MeCN gradient). The product (**8**) was purified as a white solid (23 mg, 37% yield over 2 steps). ^{1}H NMR (600 MHz, MeOD) δ 0.96 (m, 6H), 1.49 (s, 9H), 1.55 (s, 9H), 1.60 (m, 2H), 1.72 (m, 2H), 1.95 (m, 2H), 2.07 (m, 1H), 3.02 (m, 4H), 3.14−3.23 (m, 2H), 3.24-3.35 (m, 2H) 3.47 (m, 2H), 4.08 (d, J = 7.2 Hz, 1H), 4.17 (t, J = 7.5 Hz, 1H), 4.26 (m, 1H), 4.35 (m, 1H), 4.42 (m, 2H), 7.23 (m, 2H), 7.30 (m, 5H), 7.40 (m, 2H), 7.58 (t, J = 6.9 Hz, 2H), 7.81 (d, J = 7.8 Hz, 2H); HRMS (EI) *m/z* $[M + H]^+$ Calculated for $C_{47}H_{66}N_7O_8$ 856.497, observed 856.499.

Phevamine A (9)



To a solution of Fmoc-(S)-Phe-(S)-Val-spermidine-*N,N'*-Di-Boc-1*H*-guanidine (**8**) (23 mg, 0.027 mmol) in 2.5 mL of MeCN, diethylamine (300µL) was added. The reaction was stirred for 6 hours at room temperature. The crude reaction mixture was concentrated under reduced pressure. Ethyl acetate was added, and the reaction mixture was concentrated again. The reaction mixture was dissolved in CH_2Cl_2 (3 mL) on ice and trifluoroacetic acid (350 µL) was added dropwise. The reaction was stirred at room temperature for 12 hours. The reaction mixture was concentrated under reduced pressure. Toluene was added and the reaction mixture was concentrated until dryness. The reaction mixture was purified by reverse-phase HPLC (Phenomenex Kinetex C18, H_2O/MeCN gradient). The final product phevamine A (**9**) was obtained as a white solid (9.8mg, 67% yield). ^{1}H NMR (400 MHz, D_2O) δ 0.79 (d, J = 6.8 Hz, 3H), 0.82 (d *J* = 6.8 Hz, 3H), 1.49 (m, 2H), 1.60 (m, 2H), 1.84 (m, 1H), 1.87 (m, 2H), 2.97−3.20 (m, 8H), 3.20 (t, *J* = 6.8 Hz, 2H), 3.87 (d, *J* = 8.4 Hz, 1H), 4.22 (t, *J* = 7.2 Hz, 1H), 7.15 (m, 2H), 7.29 (m, 3H); HRMS (EI) *m/z* $[M + H]^+$ Calculated for $C_{22}H_{40}N_7O_2$ 434.324, observed 434.323. ^{19}F NMR (400 MHz, D_2O) δ −75.67 (s).

2.4.21 Ornithine acetyl transferase inhibition by mangotoxin

Plant ornithine acetyl transferase (OAT) was obtained from 72 g of *Arabidopsis thaliana* leaves. *A. thaliana* leaves were frozen with liquid nitrogen and crushed using a mortar and pestle until a fine powder was obtained. Leaves were then soaked for 1 hour in 200 mL plant extraction buffer consisting of 50 mM Tris base (pH 7.5), 0.2 mM cobalt chloride, 1 mM reduced glutathione, and 0.1 mM phenylmethylsulfonyl fluoride. The resulting mixture was filtered using a Buchner funnel filter and was stored –20 °C until use.

The full OAT assay consisted of 1.2 mM *N*-acetyl-L-Orn, 1.2 mM L-Glu, 50 µL plant extract containing OAT, and 25 µL of extract from 100 mL of UMAF culture. The following controls were conducted: no bacterial extract, no plant extract, no *N*-acetyl-L-Orn, no L-Glu, full assay with the extract of *Pto* extract instead of UMAF, and a control without plant extract but with *Pto* extract instead of UMAF. Assays were incubated at 37 °C for one hour, then quenched with 250 µL of 0.4 M citric acid and boiled at 95 °C for 90 seconds. After quenching, reactions were analyzed by LC-MS.

2.4.22 E. coli growth inhibition by P. syringae pv. syringae UMAF0158

E. coli growth inhibition assays were conducted using PMS agar plates using *E. coli* BL21 as the indicator strain in the top agar. The PMS media consisted of ammonium phosphate monobasic (1 g/L; pH 7), potassium chloride (0.2 g/L), magnesium sulfate (0.2 g/L), dextrose (2 g/L), and agar (1.2% w/v). Top agar contained 0.1% *E. coli* BL21 by volume. Some plates also contained L-ornithine in the top agar (30 mM). Each plate was divided into four quadrants, and 5 µL of water or King's B media was added to two quadrants as negative controls. Varying amounts (2-10 µL) of either UMAF or *Pto* overnight seed culture were added to the remaining

quadrants. Each plate was duplicated, one containing Orn in the top agar and one without Orn for reversion of *E. coli* growth inhibition by UMAF by Orn supplementation. Plates were incubated at room temperature for one day, then at 37 °C for a second day.

2.4.23 ATP-grasp sequence similarity network formation

Amino acid sequences of specified ATP-grasp enzymes were obtained from NCBI. Sequences were uploaded as FASTA files to the Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST) web-tool with an E-value of 10^{-100} unless otherwise specified.[100] When indicated, all other ATP-grasp enzymes were included by specifying the ATP-grasp Pfam value: IPR011761. Sequences were limited to only those 150-700 AA in length. Networks with nodes depicting proteins with ≥50% sequence identity were visualized in Cytoscape with alignment scores of 100.[101-102]

CHAPTER 3: TOWARD IDENTIFYING A SMALL MOLECULE PRODUCED BY A CONSERVED GENE CLUSTER FROM A HYPERVIRULENT KIWIFRUIT PATHOGEN

3.1 INTRODUCTION

Bacterial plant pathogens produce a wide range of small molecules. Small molecules, such as siderophores, phytotoxins, and antibiotics, play key roles in the fitness of the producing pathogens in the presence of plants and other microbes. *Pseudomonas syringae* infects a wide range of plants and consists of more than 50 pathovars defined based on the plant infected. Therefore, it is widely used as a model to study virulence, the ability to infect plants. *P. syringae* produces several well-known toxins: tabtoxin, coronatine, phaseolotoxin, and syringomycin, some of which inhibit plant or microbial enzymes at nanomolar concentrations.[39, 103] Recently, our group used a genome-guided approach to identify a new small molecule, phevamine A, which promotes the virulence of *P. syringae* pv. *tomato* DC3000 by suppressing plant immunity.[12]

P. syringae pv. *actinidiae* (*Psa*) is a pathogenic bacterium that leads to virulence in kiwifruit, having a devastating impact on the agriculture and economy (estimated loss of $310-410 million in the next five years and $885 million in the next 15 years in New Zealand alone).[104] Outbreaks of this bacterial strain also affect Japan, Italy, and Korea, becoming a global problem.[72] Therefore, it is essential to understand the molecular mechanisms of infection so methods to overcome and prevent infection can be devised. Here, we focus on an uncharacterized biosynthetic gene cluster conserved in many hypervirulent *Psa* strains and set out to identify the functions of the enzymes encoded in this gene cluster.

The gene cluster (denoted hereafter as *kwf*) was identified by large-scale genome sequencing and bioinformatics analysis.[72] This gene cluster was present in many virulent strains but not present in low-virulent strains, suggesting this cluster may be involved in virulence. The *kwf* cluster encodes two nonribosomal peptide synthetases (NRPSs), which form peptide bonds independently of the ribosome and are a major enzyme family for the biosynthesis of small molecules.[105] The functions of enzymes encoded by this cluster are uncharacterized. We propose that the *kwf*-encoded enzymes synthesize bioactive small molecules and focus on the cluster from *Psa* NZV-13.

NRPS genes are modular and each module consists of domains, the minimum usually being a condensation domain (C), an adenylation domain (A), and a thiolation domain (T).[106] The T domain contains a free thiol formed by transfer of 4'-phosphopantetheine from coenzyme A to an active site serine and this thiol covalently binds acyl reaction intermediates.[106] The phosphopantetheine arm is a long, flexible linkage that facilitates interaction with other NRPS domains.[106] The A domain confers substrate specificity and activates a single amino acid by adenylation, releasing pyrophosphate (PP_i).[106] The C domain then catalyzes the formation of amide bonds between amino acids loaded on the T domains.[106] Other domains commonly found in NRPSs include a reductase (R) domain, which can reduce the final product from a carboxylic acid to an alcohol or aldehyde,[106] and a methyltransferase (MT) domain may also be present, which facilitates transfer of a methyl group from S-adenosyl-L-methionine.[106]

This chapter discusses the efforts to identify the small molecule produced by *kwf*-encoded enzymes. Concurrent with structure elucidation of the small molecule, the *kwf*-encoded enzymes are investigated to understand the chemistry involved in biosynthesis. The

workflow that has been set up for the *hsv* operon is applied to the *kwf* cluster, in which metabolomics and enzymology are combined to piece together the structure of the bioactive small molecule. Here, we present preliminary results from in vitro enzymology and metabolomics for the *kwf* cluster in *Psa* NZV-13. Two substrates were identified for NRPS enzymes. Metabolomics experiments have been conducted in two different systems, *P. syringae* and *E. coli*. Results from these experiments are complementary and will aid the identification and structural elucidation of the final small molecule.

3.2 RESULTS

3.2.1 In vitro enzyme reconstitution of the kwf cluster.

Enzymology was used as a starting point for characterization of the metabolite(s) produced by the *kwf* gene cluster from *Psa* NZV-13 (Figure 3.1). Identification of preferred substrates for some enzymes encoded by the cluster can facilitate the identification of the small molecule product(s) by metabolomics. The two NRPS proteins OrfB and OrfE likely build the peptide backbone of the small molecule, and therefore were the initial focus of in vitro reconstitution. OrfB and OrfE contain a total of three A domains (Table 3.1), thus we propose that these enzymes activate and condense three amino acids.

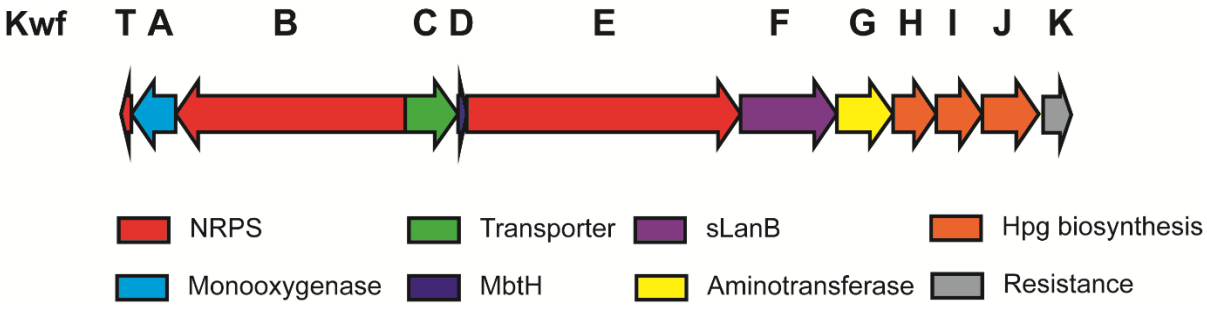


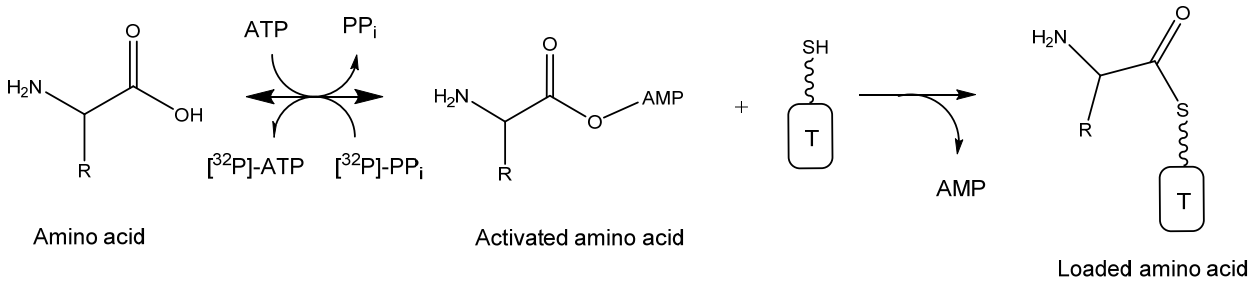
Figure 3.1. Graphical representation of the *kwf* operon.

Table 3.1. Bioinformatic analysis of the *kwf* cluster.

Protein	Predicted function(s)
OrfT	Standalone T-domain
OrfA	Luciferase-like monooxygenase
OrfB	NRPS (C-A-MT-T-R)
OrfC	MFS transporter
OrfD	MbtH-like protein
OrfE	NRPS (C-A-T-C-A-T)
OrfF	Small LanB
OrfG	Aminotransferase
OrfH	4-hydroxyphenylpyruvate dioxygenase
OrfI	α-hydroxy-acid oxidase
OrfJ	(S)-3,5-dihydroxyphenylglycine transaminase, transcriptional regulator
OrfK	Resistance

Various constructs were generated for each NRPS module. The constructs examined include full length OrfB, full length OrfE, OrfB A-MT-T, OrfB A, OrfE1 A-T-C, OrfE2 A-T, OrfE1 A, and OrfE2 A. Substrate specificity of each A domain was examined using an ATP-$^{32}PP_i$ exchange assay. Briefly, this assay involves incubating the A domain-containing enzyme with a potential amino acid substrate, the cofactors ATP and Mg^{2+}, and an excess of radioactive $^{32}PP_i$ ("hot") and regular PP_i ("cold"). If the A domain activates the amino acid present by

adenylation, the reverse reaction incorporates radioactive $^{32}PP_i$ into ^{32}ATP (Scheme 3.1). Labeled ^{32}ATP is then trapped by activated charcoal, resuspended in a scintillation cocktail, and counted using a scintillation counter. As a general rule of thumb, we used greater than 9% substrate conversion or 30 times of the counts from a no-substrate control as a threshold to consider the adenylating activity as significant.



Scheme 3.1. Reaction scheme for ATP-$^{32}PP_i$ exchange assays. The NRPS A domain confers specificity for a specific amino acid. This amino acid is then acylated by ATP, releasing pyrophosphate (PP_i) and forming an aminoacyl bond with AMP. This amino acid can then be loaded onto the T domain, forming a thioester linkage. Radioactive $^{32}PP_i$ can be added to the A domain reaction, and the reverse reaction will incorporate ^{32}P into ATP, which can be adsorbed to activated charcoal and the radioactivity can be measured by scintillation counting.

NRPS constructs were first tested with amino acid substrates that had been predicted by NRPS prediction online software.[107-108] OrfB is predicted to activate hydroxyphenylglycine, which is likely synthesized by the OrfH, OrfI, and OrfJ proteins encoded by the *kwf* cluster. These three proteins are homologous to HmaS, Hmo, and HpgT, respectively, which synthesize hydroxyphenylglycine (hpg) in *Amycolatopsis orientalis* (48-50% sequence identity) and incorporate it into the natural product chloroeremomycin.[109] In contrast to this prediction, none of the OrfB constructs activated D- or L-hpg under our assay conditions, but instead, OrfB A activated L-Phe with a 40% overall conversion (Fig. 3.2). Further, the MbtH protein OrfD was necessary for this activity; in the absence of OrfD, the activity observed is

0.8% of that of the reaction in the presence of OrfD (Fig. 3.2). The first module of OrfE, OrfE1 A-T-C, was predicted to activate L-Ser, and we found it was able to activate both L-Ser and L-Ala with 47% and 41% conversion, respectively (Fig. 3.3). We were unable to reconstitute the activity for the second OrfE module, in which the 10 amino acid code of the A domain only has weak bioinformatic prediction, such as L-His, L-Glu, or β-Ala. All 20 proteinogenic amino acids (excluding Trp) and β-Ala were tested, but no activation was observed (Fig. 3.4). Other tailoring enzymes are also present in the *kwf* cluster with putative functions assigned through bioinformatics analyses, but these enzymes have not been the focus of current in vitro enzymatic assays (Table 3.1, Fig. 3.1).

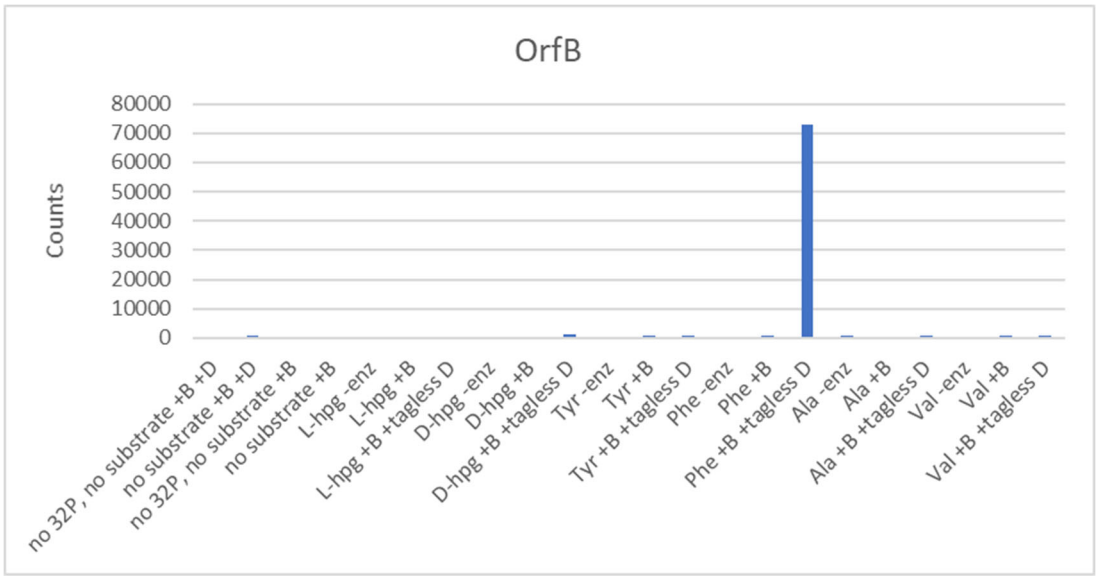


Figure 3.2. OrfB prefers to activate l-Phe based on results from ATP-$^{32}PP_i$ exchange assays. The A domain of OrfB selectively activates L-Phe instead of D- and L-hydroxyphenylglycine.

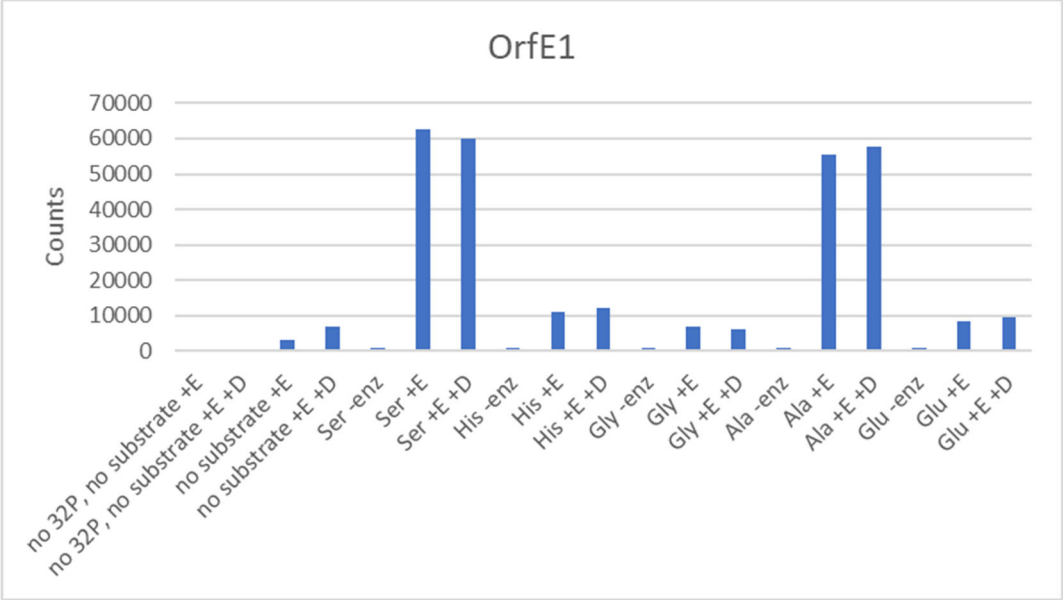


Figure 3.3. OrfE1 prefers to activate l-Ser and l-Ala based on results from ATP-$^{32}PP_i$ exchange assays. The A domain of the first OrfE module activates both L-Ser and L-Ala.

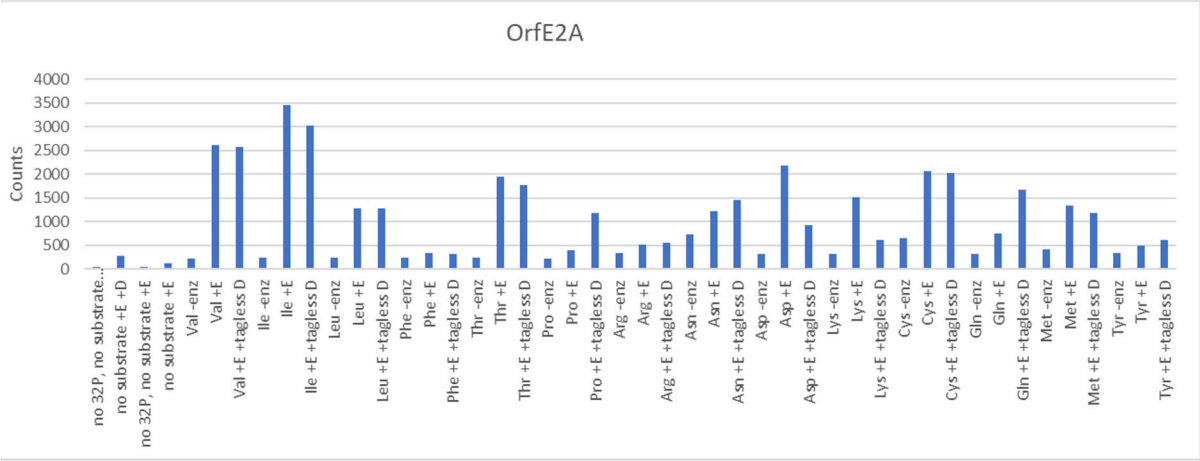


Figure 3.4. OrfE2 does not prefer any canonical amino acids based on results from ATP-$^{32}PP_i$ exchange assays. The A domain of the second OrfE module does not activate any of the 19 proteinogenic amino acids tested or β-Ala.

3.2.2 Metabolomics analysis of Psa NZV-13 and cloning and heterologous expression of the kwf cluster in E. coli.

NRPS deletion mutants of *orf*B and *orf*E were generated by Paul Rainey's group in New Zealand, including single deletion mutants of each gene and a double deletion mutant.

Each strain was grown in triplicate alongside wildtype NZV-13 and all culture supernatants were extracted with chloroform-methanol. Cultures were grown in both M9 minimal media and KB rich media. Both solvent layers were dried down and analyzed by LC-HRMS, and the metabolomic profiling data was uploaded to XCMS. Features were sorted by fold change. Several mass-to-charge ratios (*m/z*) have been identified that are significantly more abundant in wildtype extracts than any of the knockout extracts. We expect the product of the *kwf* cluster to contain a tripeptide backbone synthesized by the NRPS enzymes, and therefore we expect the mass of the product to range from 400-700 Da, so we looked for masses in that range to identify the final small molecule product as well as below that range to identify intermediates. Some mass features were identified that fit the above criteria and have the following *m/z* values (and fold changes between wildtype and double mutant): 254.1388 (20,557), 324.1376 (5,966), 320.1348 (5,917), 508.2406 (3,754), and 336.1232 (3,244).

Metabolomics experiments were also conducted using a heterologous expression system. The entire 22.5 kb *kwf* cluster was cloned into the pETcoco-1 plasmid using Gibson assembly, then *E. coli* Bap1 was transformed with either the full construct or empty pETcoco-1 (Fig. 3.5). Since *orfB*, *orfA*, and *orfT* (BAT) face the opposite way of the rest of the cluster (Fig. 3.1), BAT was cloned after *orfK* with its own T7 promoter to ensure all genes would be transcribed. Therefore, a T7 promoter was present in the plasmid before *orf*C that promotes transcription through *orfK*, and then a second T7 promoter was present before *orfB* to promote transcription through *orfT* (Fig. 3.5). This construct was expressed in *E. coli* and compared to empty pETcoco-1 using the same technique as above for *Psa* NVZ-13 and knockout strains. Again, mass features with *m/z* values in the 400–700 range were selected. Molecules with the highest fold changes were the following: 719.2382 (233 fold change), 604.2273 (139),

443.1278 (118), 436.1342 (72), 626.2217 (58), 452.1177 (55), 606.2371 (51), 699.3995 (32), and 720.3989 (18). Two molecules were identified in both *P. syringae* and *E. coli* metabolomics with high fold changes and the m/z values of these molecules were 231.1714 and 137.0479. The masses of these molecules are smaller than expected, suggesting the molecules are biosynthetic intermediates or shunt products due to overexpression of the *kwf* cluster in *E. coli*.

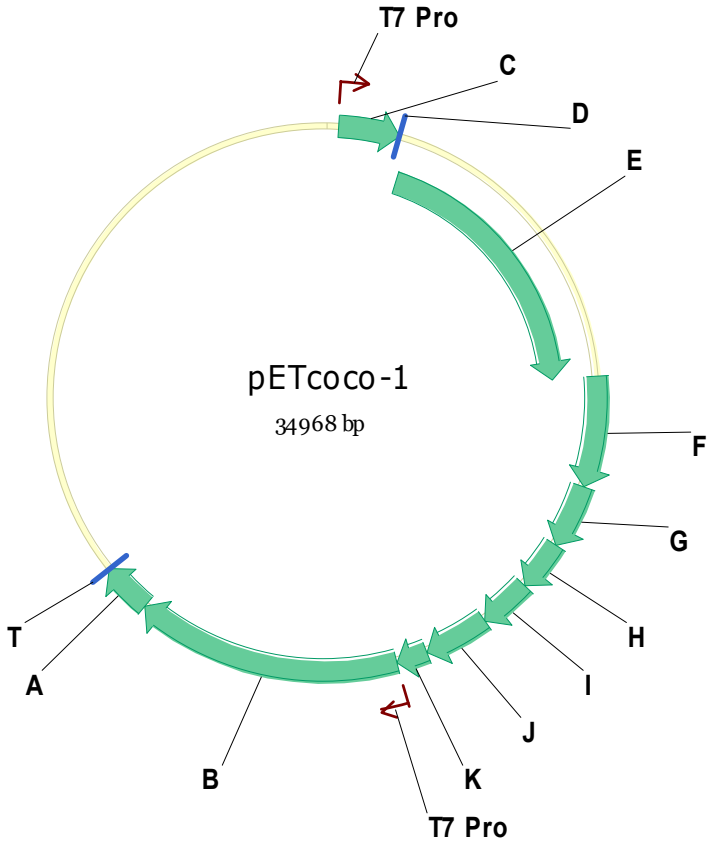
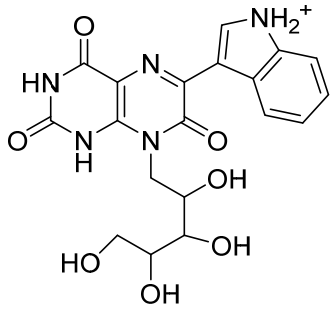


Figure 3.5. Schematic of the kwf cluster cloned in pETcoco-1. The entire construct is 35 kb and contains a T7 promoter before *orfC* and another before *orfB*.

3.2.3 Antibacterial activity testing of small molecules isolated from Psa NZV-13.

Psa NZV-13 was found to produce a small molecule with a mass-to-charge ratio of 430.1357. This molecule was named KB-1 and the chemical structure was elucidated by a

postdoc in the group, Wenlong Cai (Fig. 3.6). Stereochemistry has not yet been characterized. Purified compound was provided by Dr. Cai and antibacterial assays were conducted using the agar diffusion method. No activity was observed against any strain tested as a target. Subsequently, whole cell culture of the wildtype strain was used for antibacterial assays. Again, no activity was observed against any strain tested. The KB-1 purified compound and whole cell NZV-13 culture were tested against the following strains: *P. syringae* pv. *syringae* UMAF0158, *P. syringae* pv. *syringae* B728a, *P. syringae* pv. *tomato* DC3000, *P. syringae* pv. *pisi*, *P. syringae* Cit7, *P. fluorescens* Pf0-1, *P. fluorescens* Pf0-5, *P. putida* S12, *P. putida* F1, *P. aeruginosa* PA14, *E. coli* MG1655, *Micrococcus luteus*, *Enterococcus faecalis*, and *Bacillus subtilis* BSU168. Select representative strains tested against KB-1 are shown in Figure S3.1. KB-1 was also tested against the fungus *Rhizoctonia solani*, but no inhibition was observed against fungal growth.



Exact Mass: 430.1357

Figure 3.6. Chemical structure of KB-1. KB-1 was isolated from *Psa* NZV-13 culture and the structure was assigned by Wenlong Cai using one- and two-dimensional NMRs. Stereochemistry has not been characterized.

3.3 DISCUSSION

3.3.1 In vitro enzyme reconstitution of the kwf cluster.

OrfB was found to selectively activate L-Phe and the first module of OrfE was found to activate L-Ser. OrfB activity was only observed when the MbtH protein OrfD was present, but the activity of OrfE1 did not require OrfD. MbtH proteins are thought to act as chaperone proteins to bind to an A domain and induce enzyme activity of the A domain.[110-111] OrfB was predicted to activate hpg. This prediction was further supported by the homology between OrfH, OrfI, and OrfJ and the hpg biosynthetic enzymes. However, our results suggest that OrfB activates L-Phe instead of hpg. The homologs to hpg enzymes in the *kwf* cluster may catalyze similar chemistry to hpg biosynthesis but synthesize phenylglycine. Another interesting observation is that only certain NRPS constructs yielded activity for ATP-$^{32}PP_i$ exchange assays. The OrfB-A-MT-T construct didn't activate L-Phe, but OrfB-A did. Similarly, OrfE1-A-T-C activated L-Ser and L-Ala, but OrfE1-C-A-T and OrfE1-A did not activate any amino acids. This observation indicates it is necessary to generate several constructs for each NRPS module to identify a properly folded and active construct. To determine whether L-Ser or L-Ala is the native substrate for OrfE1, kinetics experiments will be performed to obtain Michaelis-Menten parameters. If the measured k_{cat} and K_M exist in a range relevant under physiological conditions, the amino acid is likely the native substrate. Lastly, since no activity was observed for OrfE2-A, OrfE2-C-A-T, or OrfE2-A-T, this enzyme might activate a noncanonical substrate, or these constructs may not be active.

Identification of substrates accepted by other *kwf*-encoded enzymes and the chemistry these enzymes catalyze will help piece together the small molecule product. Due to the lack of any thioesterase NRPS domain, which could cyclize the peptide product, the amino acid

substrates will most likely be linked together as a linear peptide and not a cyclic peptide. An *N*-methyltransferase domain is also present in OrfB, which likely methylates the nitrogen of phenylalanine. OrfF is a putative small LanB (sLanB) enzyme, which contains only one of the two domains in LanB enzymes. LanB enzymes glutamylate an amino acid, typically serine or threonine, and then eliminate the glutamate to form dehydroalanine or dehydrobutyrine [112]. Small LanB enzymes only possess the first domain, so they likely catalyze glutamylation of an amino acid sidechain but not dehydration. OrfG shares sequence homology with pyridoxal phosphate-dependent aminotransferases that catalyze the conversion between an alpha-keto acid and an amino acid. OrfC is annotated as a membrane transport protein, potentially transporting the small molecule out of (or into) a cell. The remaining protein is OrfA, which shares 48% sequence identity with luciferase-like monooxygenases, so it could be incorporating oxygen into the final product via hydroxylation, epoxidation, or Baeyer-Villiger oxidation.[113] OrfB contains a C-terminal reductase domain, suggesting that the NRPS product is released as an aldehyde or an alcohol, and OrfB contains the termination module. Thus, the action of OrfE likely precedes that of OrfB. OrfE contains an N-terminal C domain, which likely acts as a starter C domain that links a lipid or another moiety to the N-terminus of the molecule. Reconstitution of the activities of these enzymes requires knowledge of their potential substrates and will be accelerated once the small molecule produced by the *kwf* cluster is identified.

3.3.2 Metabolomics analysis of Psa NZV-13 and cloning and heterologous expression of the kwf cluster in E. coli.

We expect the product of the *kwf* cluster to be a tri- or tetrapeptide due to the following rationale. The NRPS enzymes contain a total of three A domains which can activate three

amino acid substrates, but there are three condensation domains to link these, so there may be three amide bonds and therefore four amino acids. Interestingly, there is a standalone T domain upstream of *orfA*, but no corresponding A domain, so it is unclear what substrate may be tethered there natively. Further, OrfF, the sLanB enzyme, may be glutamylating the compound and therefore adding one more amino acid. Therefore, the expected product mass could range anywhere from ~400-700 Da. From enzymatic data, we expect the compound to contain a serine and a phenylalanine/phenylglycine (pg) due to OrfE1 and OrfB activity, respectively. OrfE2 may be activating a noncanonical residue and OrfF may be adding a glutamate residue to the molecule. Several potential structures were proposed based on these stipulations as a starting point. Further modifications were then also taken into account including the aminotransferase OrfG, the methyltransferase domain of OrfB, and the putative monooxygenase OrfA. From aforementioned *P. syringae* metabolomics data, the metabolite with an *m/z* of 324.1376 could correspond to the Phe-Ser-Ala tripeptide, and the metabolite with an *m/z* of 336.1232 could correspond to pg-Ser-Pro. However, these structures are likely to be intermediates of the *kwf* pathway and none of the other hypothesized structures match with the metabolite features (*m/z*) identified from XCMS analysis. Further in vitro studies will need to be conducted before metabolomics data can be analyzed further.

3.3.3 Antibacterial activity of Psa NZV-13 culture and isolated natural product.

Neither the compound KB-1 nor the wildtype *Psa* NZV-13 possess antibacterial activity against any strain tested. *Bacillus subtilis* is a representative Gram-positive strain of bacteria and the *E. coli* strain tested is a representative Gram-negative strain. Further, the *E. coli* strain used has a weakened cell membrane (imp), making it more susceptible to

antibacterial compounds, indicating that NZV-13 is not a potent antibiotic against *E. coli*. Several other species and strains of *Pseudomonas* were tested as well since pseudomonads are soil bacteria and interactions and competition among different pseudomonads likely occur in the soil. However, no antibacterial activity was observed against the pseudomonads tested. Similarly, activity was not observed against the representative *Enterococcus* or *Micrococcus* strains, both of which are Gram-positive bacteria. Therefore, it seems that NZV-13 does not produce antibiotics under the laboratory conditions tested, at least not in high enough quantities to be detected in agar diffusion assays. Other growth conditions will need to be tested to further explore the antimicrobial activities of *Psa* NZV-13.

3.4 CONCLUSION AND OUTLOOK

Several experiments are currently underway to characterize the product of the *kwf* cluster and its biosynthesis. The following enzyme constructs have been cloned, expressed, and are ready to be purified for in vitro enzyme assays: OrfT, OrfA, and OrfG. Purified protein has already been obtained for OrfF, OrfH, OrfI, and OrfJ. Both hydroxyphenylpyruvate and phenylpyruvate will be incubated with OrfH, OrfI, and OrfJ consecutively to synthesize hpg or pg, respectively. Other purified constructs of OrfB will be tested for the methyltransfer reaction catalyzed by the MT domain and reduction to an alcohol or aldehyde catalyzed by the reductase domain. OrfB has already been purified as the full C-A-MT-T-R construct and as A-MT-T for these experiments. Isotope-enriched L-Phe will also be used for metabolomics experiments, based on in vitro activation data by OrfB, which can help narrow down the data output from XCMS. OrfE2 A will be tested with noncanonical NRPS substrates while other OrfE2 constructs are generated and purified for testing. Once in vitro experiments and

metabolomics experiments suggest a common mass, the molecule will be purified for subsequent structure elucidation by tandem MS and NMR. Once the natural product synthesized by the *kwf* cluster is purified and characterized, it will be subjected to biological assays with plants to characterize its mechanism of action.

In this project, the same approach was used as the phevamine A project with the goal of connecting uncharacterized biosynthetic gene clusters to bioactive small molecules. The dual approach combining metabolomics with in vitro enzymology can be widely applied to identify natural products based on biosynthetic gene clusters. These studies will likely yield new natural product structures and biosynthetic reactions. The natural products we identify from phytopathogens will likely advance the understanding of the virulence of these pathogens and pave the way for overcoming bacterial virulence to treat crops against various infectious diseases.

3.5 EXPERIMENTAL PROCEDURES

3.5.1 General

All chemicals were purchased from Sigma-Aldrich or Fisher. Primers were ordered from Eton Bioscience (Table S3.1). Polymerase, restriction enzymes, and ligase were purchased from New England Biolabs. DNA manipulation kits were purchased from Qiagen, Zymo, and Thermo Scientific. PCR reactions were carried out using an Eppendorf Nexus GSX1 thermocycler with Q5 DNA polymerase. General genetic manipulation of *E. coli* was carried out using standard protocols.[89] LC-MS analysis was conducted using an Agilent 6520 accurate-mass Q-TOF LC/MS.

3.5.2 Bacterial strains and culture conditions

For cloning, *E. coli* Top10 chemically competent maximum efficiency cells were used (Invitrogen). For heterologous expression, *E. coli* Bap1 or *E. coli* BL21(DE3) electrocompetent cells (Agilent) were transformed with vectors containing expression constructs. BAP1 was used instead of BL21 for constructs containing a full NRPS gene. All constructs are listed in Table S3.2.

All overnight seed cultures were inoculated with a single colony and grown in 5 mL media in 14 mL Falcon culture tubes. Antibiotics were used in the following concentrations unless otherwise noted: 25 µg/mL chloramphenicol, 50 µg/mL kanamycin, or 100 µg/mL ampicillin. All cultures for metabolomics were grown in 100 mL media with the appropriate antibiotic and 100 µL−2 mL of a saturated overnight culture. All *E. coli* cultures were grown in LB Media (Miller, granulated, Fisher) at 37 °C to an OD_{600} between 0.4 and 0.8. *E. coli* cultures were induced with 0.5 mM IPTG and then were incubated at 16 °C overnight, shaking at 225 rpm. *P. syringae* cultures were grown overnight in King's B Media (20 g/L Bacto Proteose Peptone #3, 1% glycerol by volume, 1.5 g/L K_2HPO_4, and 0.734 g/L $MgSO_4$, dissolved in water). When specified, some *P. syringae* cultures were resuspended in M9 once reaching an OD_{600} of >0.6 in King's B media. In this case, *P. syringae* cultures were washed with 100 mM $MgCl_2$ and resuspended in M9 minimal media (12.8 g/L Na_2HPO_4 ● 7 H_2O, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L NH_4Cl, 1% glycerol by volume, 0.024 g/L $MgSO_4$, 0.0011 g/L $CaCl_2$, and 0.18 g/L mannitol, dissolved in water). *P. syringae* cultures were grown at 28 °C overnight and 16 °C for another 24 hours while shaking at 225 rpm and were not induced. When *P. syringae* was supplemented with phevamine A, the molecule was added once the cells were resuspended in minimal media to a final concentration of 5 µM. When spent media was

added to *E. coli* heterologous expression cultures, a 2-day-old 100 mL *Psa* NZV-13 culture was gently pelleted and 10 mL supernatant was added to 90 mL LB containing chloramphenicol and 1 mL *E. coli* harboring the *kwf* cluster or empty pETcoco-1 vector control. These cultures were grown at 37 °C and induced as usual.

3.5.3 Gibson assembly of the kwf cluster

orfBAT was first cloned into pET30a and then amplified for Gibson to incorporate a T7 promoter into the fragment. pETcoco-1 was digested with HindIII-HF (NEB) and blunt ends were generated using mung bean nuclease (NEB) to prevent plasmid recircularization. Three fragments (CDE, FGHIJK, and T7-BAT) were amplified by PCR and assembled before adding the digested pETcoco-1 backbone to ensure BAT didn't anneal to the T7 promoter on the plasmid. Each fragment was added in equal amounts (300 ng each) with 2X HiFi DNA Assembly Mix (NEB) to a total volume of 26 µL. The reaction was incubated 50 °C for 20 min, then 600 ng of digested pETcoco-1 was added with an equal volume of additional 2X HiFi DNA Assembly Mix and the reaction was incubated at 50 °C for another 1 hour.

3.5.4 Metabolite extraction for metabolomics

E. coli and *P. syringae* cells were removed by centrifugation of the cell culture (100 mL) at 3500 x *g* and 4 °C for 10 minutes. Solvent extractions of the culture supernatant were performed using an equal volume of supernatant to chloroform-methanol (2:1:1). The top aqueous layer was separated into a round-bottom flask and concentrated under vacuum. The dried down material was resuspended and transferred to a small glass vial, concentrated under reduced pressure, and stored at −20 °C until MS analysis.

3.5.5 Protein expression and purification

Large-scale cultures for protein purification were grown in 1 L of media containing the appropriate antibiotic and 1−5 mL of overnight culture. At an OD_{600} of 0.5−0.8, a sample of 0.5 mM IPTG was added to induce protein expression. After 16 hours of protein expression, cell pellets were collected by centrifugation and resuspended in phosphate lysis buffer (50 mM K_2HPO_4, 150 mM NaCl, 10% glycerol, pH 8.1). Cells were lysed by sonication at 30% amplitude for 3 minutes (0.5 seconds on and 1.5 seconds off). Cell debris was removed by centrifugation at 17,000 rpm for 40 minutes and the supernatant was filtered using a 0.45 μm filter (Corning). Each Orf protein (excluding OrfD) was purified using an AKTA FPLC using a 5 mL HiTrap nickel column (GE) at a flow rate of 3.0 mL/min. Wash buffer consisted of 50 mM K_2HPO_4, 150 mM NaCl, 25 mM imidazole, and 10% glycerol (pH 8.1). Elution buffer was the wash buffer with the addition of 250 mM imidazole. Tagless OrfD was purified using a 5 mL HiTrap Q HP anion exchange column (GE) at a flow rate of 3.0 mL/min. Wash buffer consisted of 20 mM TRIS (pH 8.0). Elution buffer consisted of 20 mM TRIS (pH 8.0) and 1 M NaCl. The wash step lasted for 5 column volumes (CVs). All proteins were eluted using a gradient of 0−100% of elution buffer over 10 CVs, and then held at 100% for 5 CVs. All Orf proteins were further purified using a HiLoad 16/600 Superdex 200 prep grade size exclusion column (GE) on the FPLC using the lysis buffer described above. The column was pre-equilibrated with lysis buffer at a flow rate of 0.5 mL/min for 1.5 CVs, then protein was loaded onto the column and eluted over 1.5 CVs using the same buffer and flow rate.

3.5.6 MS and MS/MS analysis

Culture extracts and in vitro reactions were separated on a Phenomenex Kinetex 5 µm C18 column (100 Å, 150 mm x 4.60 mm) and analyzed using Agilent 6520 accurate-mass Q-TOF LC/MS. Mobile phases were water and acetonitrile, each containing 0.1% formic acid. The flow rate was 0.4 mL/min. The gradient was held at 2% acetonitrile for 5 minutes before ramping to 20% acetonitrile over 10 minutes, then to 65% over 5 minutes, then 95% over 10 minutes, and held at 95% for 5 minutes.

3.5.7 Agar diffusion bacterial growth inhibition assays

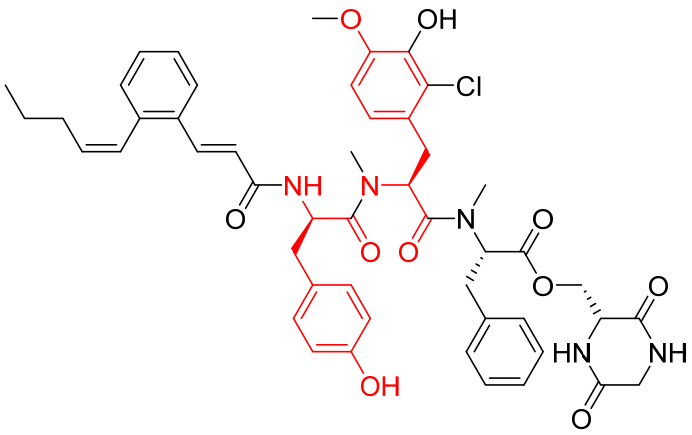
Bacterial growth inhibition assays were conducted on M9 or LB agar plates. LB media consisted of tryptone (10 g/L), sodium chloride (10 g/L), yeast extract (5 g/L), and agar (1.2% w/v). M9 media consisted of sodium phosphate dibasic heptahydrate (12.8 g/L), potassium phosphate monobasic (3 g/L), sodium chloride (0.5 g/L), ammonium chloride (1 g/L), glycerol (1% v/v), magnesium sulfate (0.024 g/L), calcium chloride (0.0011 g/L), mannitol (0.18 g/L), and agar (1.2% w/v). An overnight seed culture of each bacterium tested was diluted 1:1000 with water and a sample of 100 µL was plated and allowed to dry. One filter paper disc was placed into each quadrant of a plate. One quadrant contained a water negative control and KB-1 was added in three varying concentrations on each plate. Alternatively, a sample of 10 µL of *Psa* NZV-13 overnight seed culture was added to two quadrants and compared to two Δ*orfB*Δ*orfE* quadrants. Plates were incubated at 28 °C for 24 hours, or 24 hours at 37 °C for *E. coli* indicator plates.

CHAPTER 4: ISOLATION OF PEPTICINNAMIN E AND CHARACTERIZATION OF ITS BIOSYNTHETIC GENE CLUSTER FROM *ACTINOBACTERIA BACTERIUM* OK006

4.1 INTRODUCTION

Natural products have important applications in improving human health and treating disease. For example, many natural products and derivatives exhibit potent antitumor activities, and have been and are continuing to be explored in cancer therapy. Pepticinnamins are a class of natural products that inhibit protein farnesyltransferase (PFT).[114] Approximately 30–40% of all human tumors, and more than 80% of major tumors from colon, breast, and pancreatic cancer, contain mutations in genes encoding Ras proteins.[114-115] Tumor-associated Ras requires *S*-farnesylation of its C-terminus for function in promotion of transduction of growth signals, which ultimately leads to cell proliferation and malignancy,[114-116] thus the enzymes responsible for Ras farnesylation, the PFTs, are promising anticancer targets.

Pepticinnamins were first isolated from the actinomycete *Streptomyces* sp. OH-4652.[117] Six pepticinnamin variants have been characterized, denoted pepticinnamin A-F, and are all potent inhibitors of PFT. The most potent in this class is pepticinnamin C with an IC_{50} of 100 nM, and the least potent is pepticinnamin D with an IC_{50} of 1 μM.[118] Literature studies focus on pepticinnamin E, the most abundant pepticinnamin produced by *Streptomyces* by ten-fold, which has an IC_{50} of 300 nM.[118] The six pepticinnamins have been separated by HPLC and *m/z* values are reported for all, but only pepticinnamin E has been structurally characterized (Fig. 4.1).[114-115, 117-119] The producing strain *Streptomyces* sp. OH-4652 has not been sequenced and the biosynthetic pathway for the pepticinnamins was unknown.



pepticinnamin E

Figure 4.1. Structure of pepticinnamin E. Bonds and atoms highlighted in red show the two tyrosine residues likely encoded by the NRPS Pcm2.

In this chapter, we report the first isolation and purification of pepticinnamin E from *Actinobacteria bacterium* OK006. We also identify the gene cluster responsible for production of the pepticinnamins, which contains two nonribosomal peptide synthetases and several other tailoring enzymes. Bioinformatic analysis and preliminary biochemical characterization suggest that the nonribosomal peptide synthetase (NRPS) Pcm2 likely activates and condenses two L-tyrosine residues with the assistance of the MbtH protein Pcm3. Finally, we propose that the putative hydroxylase enzyme Pcm11 catalyzes hydroxylation of L-tyrosine using a biopterin cofactor, while L-tyrosine is bound to the NRPS T domain.

4.2 RESULTS

4.2.1 Isolation of pepticinnamin E

The gene cluster encoding the pepticinnamins was identified by Kevin Santa Maria in the group through bioinformatic analyses (Fig. 4.2, Table 4.1). He searched in sequence

genome databases for biosynthetic gene clusters that encode a potential halogenase along with enzymes that could synthesize the diketopiperazine moiety and cinnamoyl tail. The genome of the *Actinobacteria bacterium* OK006 strain was identified, harboring a candidate gene cluster for pepticinnamin biosynthesis. The strain was acquired and initial metabolomics analysis showed production of a compound with the same *m/z* as pepticinnamin E. An insertional mutant of the NRPS gene *pcm2* in the cluster was generated by Andy Chan in the group, and metabolomics analysis showed that *pcm2* deletion abolished production of pepticinnamin E. This result, along with bioinformatic analysis of the gene cluster, confirmed that the *pcm* gene cluster is responsible for pepticinnamin production.

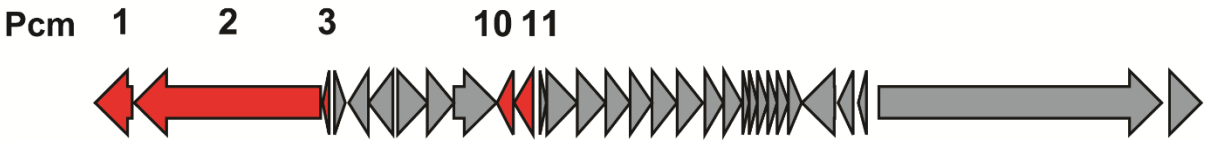


Figure 4.2. Gene cluster responsible for pepticinnamin production. Genes in red are relevant to this study.

Table 4.1. Predicted protein functions of select pepticinnamin enzymes.

Protein	**Predicted Function**
Pcm1	Halogenase
Pcm2	NRPS (C-A-T-C-A-MT-T)
Pcm3	MbtH-like protein
Pcm10	Methyltransferase
Pcm11	Tyrosine hydroxylase

To confirm the structure of pepticinnamin isolated from *A. bacterium* OK006, and to obtain materials for biological studies, we cultured the bacterium in large scale. Several liters of *A. bacterium* OK006 culture were grown. Cells were collected from the culture by gentle centrifugation and the cell pellets were extracted with methanol. Extracts were washed with water then with hexanes to remove polar and nonpolar impurities, respectively. Crude extracts were then purified by three rounds of reverse-phase preparative HPLC and then two rounds of reverse-phase analytical HPLC. Each round of purification was analyzed by LC-MS to confirm the presence pepticinnamin E. A major challenge is that in each round of purification several peaks are present that exhibit the same *m/z* of 908 as pepticinnamin E, which suggests structural isomers of pepticinnamin E are also present. The most abundant compound with the correct *m/z* value is the one pursued for further purification and structural analysis. On average, extractions yielded approximately 0.4 mg of pepticinnamin E per liter of culture. After 5 rounds of purification, proton NMR of the purified molecule identified most of the expected proton shifts for pepticinnamin E, supporting that the compound isolated from *A. bacterium* OK006 is identical to pepticinnamin E. The spectrum also contained several proton peaks we cannot account for, suggesting that some impurities are still present (Fig. 4.3, Table S4.3). Unfortunately, further purification of the material led to the loss of pepticinnamin. Further experiments are needed to obtain pure pepticinnamin E for biological studies.

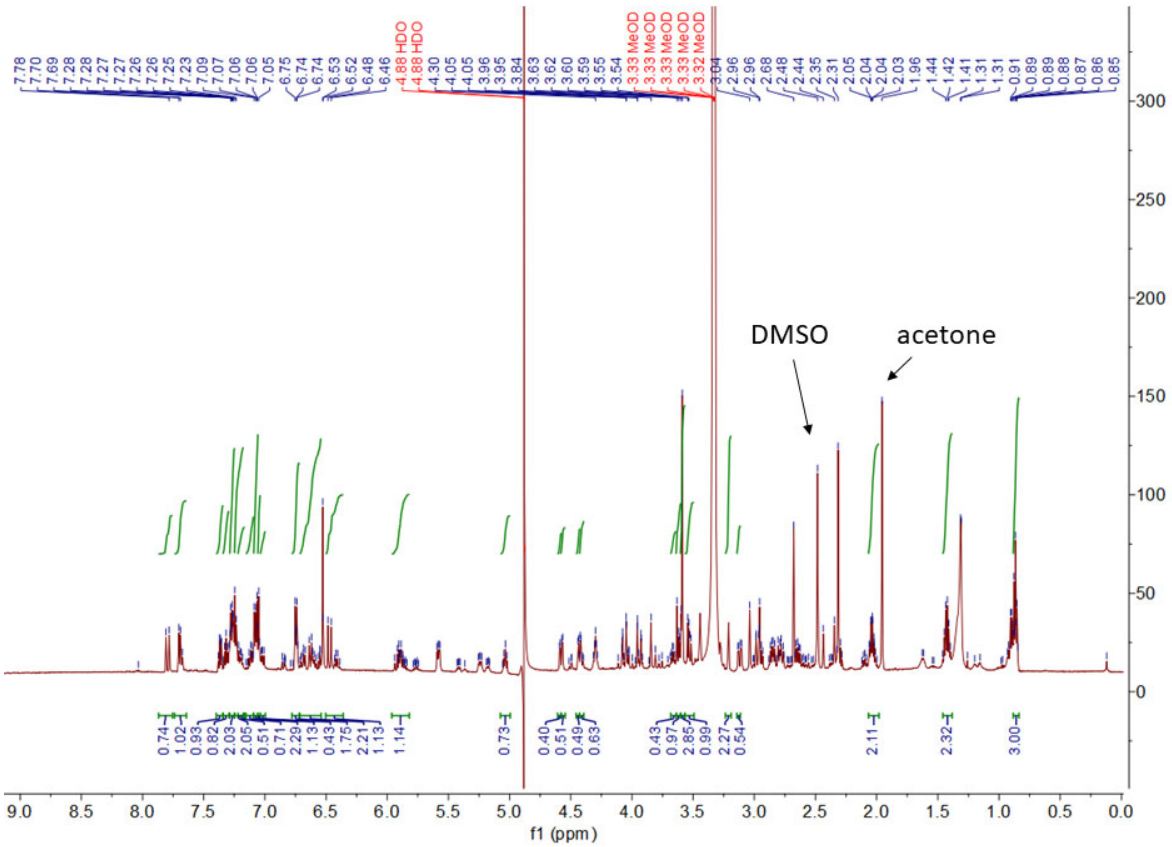


Figure 4.3. ^{1}H NMR (600 MHz) of pepticinnamin E. Spectrum acquired in MeOD. Integrated peaks correspond to signals expected for pepticinnamin E. DMSO and acetone contaminants are labeled on the spectrum, water and methanol signals are highlighted in red at the top.

4.2.2 Expression and characterization of Pcm NRPSs and accessory proteins

The NRPS Pcm2 contains two modules with the domain architecture C-A-T-C-A-MT-T. Several truncations, along with full-length protein, were designed, cloned, expressed, and purified. Each construct was inspected for its ability to activate tyrosine and similar substrates by ATP-$^{32}PP_i$ exchange assays. We expect the A domains in both modules to activate tyrosine (one D-Tyr and one L-Tyr) (Fig. 4.1), but did not observe this activity in any construct we tested.

Several assay conditions and components were tested to optimize enzyme activity. Activity of Pcm2.1 was tested at both pH 8 and pH 9, and Pcm2.1 appears to exhibit slightly more activity at the higher pH (Fig 4.4). However, the activity was very low, only 2–3 folds above the background signal. We obtained soluble proteins for full length Pcm2 coexpressed with the MbtH Pcm3. However, since we expect both modules to activate tyrosine, then if activity were to be observed by ATP-$^{32}PP_i$ exchange assays, then the activity could not be attributed to the responsible Pcm2 module. Therefore, each module of Pcm2 was individually cloned, expressed, and purified, including Pcm2.1 (C-A-T) and Pcm2.2 (C-A-MT-T). Pcm2.1 was coexpressed with Pcm3, but Pcm2.2 was not. Interestingly, Pcm2.1 was significantly more soluble than Pcm2.2, even though Pcm2.2 was expressed with an MBP tag to improve solubility. Both purified Pcm2.1 and Pcm2.2 have been tested for activity with the following substrates: L-Ala, L-Cys, L-Glu, L-Lys, L-Thr, L-Phe, L-Trp, L-Tyr, D-Tyr, and L-DOPA (Fig. 4.5, Fig. 4.6). Both modules were tested in the presence of purified Pcm3 in vitro, but no enzyme activity was observed for any substrate tested. Subsequently, both individual A domains were cloned and expressed, and the activities of these domains will be examined by ATP-$^{32}PP_i$ exchange assays as the next step.

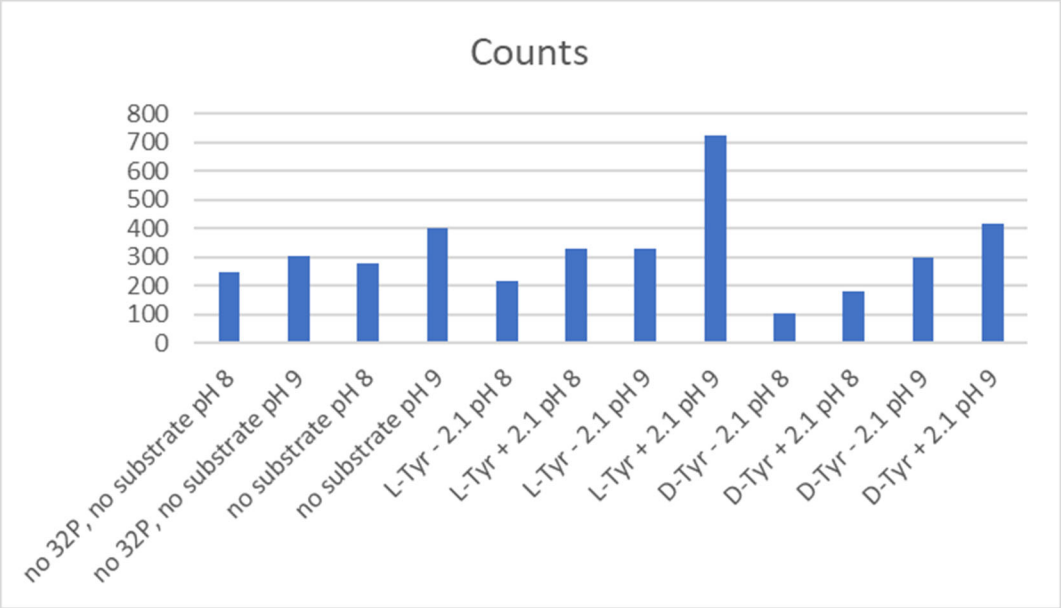


Figure 4.4. ATP-$^{32}PP_i$ exchange results for Pcm2.1 activity as a function of pH. Pcm2.1 appears to be slightly more active at pH 9 than pH 8.

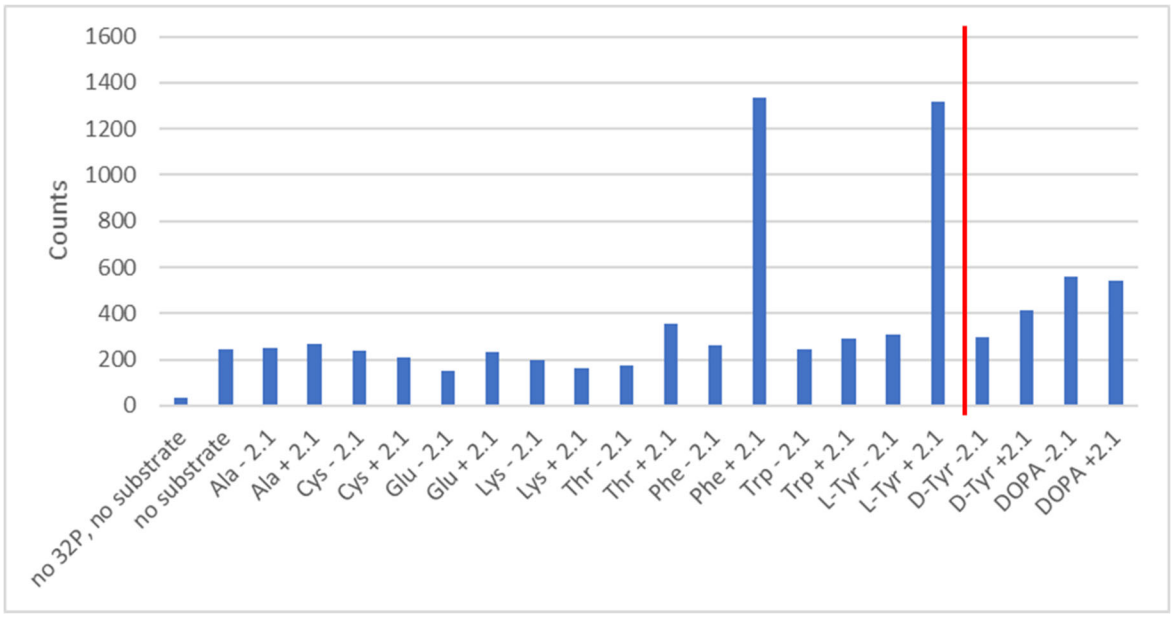


Figure 4.5. ATP-$^{32}PP_i$ exchange results for Pcm2.1. Pcm2.1 activated both L-Phe and L-Tyr at very low levels (0.7% conversion each). D-Tyr and L-DOPA assays were run in a separate experiment, as indicated by the red line, therefore direct comparisons between these substrates cannot be made.

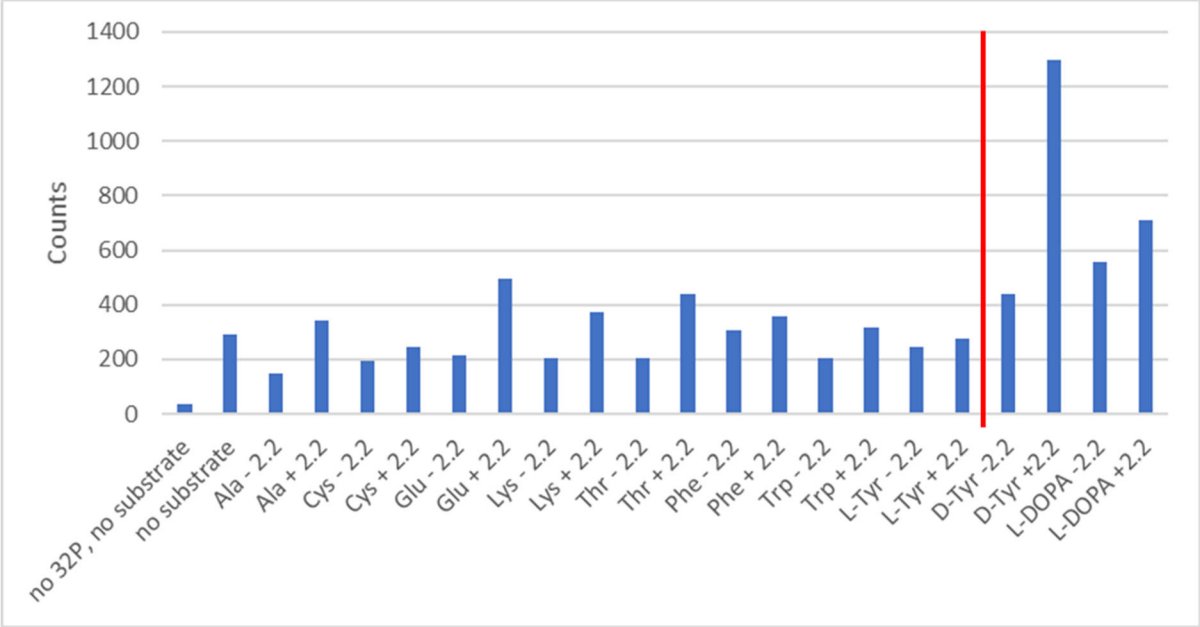


Figure 4.6. ATP-$^{32}PP_i$ exchange results for Pcm2.2. Pcm2.2 did not activate any of the substrates tested. D-Tyr and L-DOPA assays were run in a separate experiment, as indicated by the red line, therefore direct comparisons between these substrates cannot be made. Percent conversion for D-Tyr as a substrate is 0.1%.

4.2.3 Characterization of the tyrosine hydroxylase Pcm11

One tailoring enzyme encoded by the *pcm* gene cluster is the putative tyrosine hydroxylase, Pcm11. Interestingly, this enzyme is homologous to tyrosine hydroxylases in mammals and human parasites that convert L-Tyr to L-DOPA using a biopterin cofactor, but these enzymes are rare in bacteria.[120-121] One of the few bacterial examples includes the phenylalanine hydroxylase PacX from *Streptomyces coeruleoribudus*.[122] We hypothesize that Pcm11 catalyzes hydroxylation of tyrosine in the *meta* position. Kevin Santa Maria showed that the methyltransferase Pcm10 methylates 3,4-dihydroxy-L-phenylalanine (DOPA) but not L-Tyr, suggesting that the hydroxylation by Pcm11 occurs prior to methylation by Pcm10.

Timing of hydroxylation in relation to chlorination at the 2-position remains to be characterized.

Preliminary results showed that Pcm11 did not modify D- nor L-tyrosine in the presence of tetrahydrobiopterin (BH4) or (±)-6-Methyl-5,6,7,8-tetrahydropterine (6MPH4). Catalase was also present in reactions to break down peroxide byproducts of the biopterins to prevent non-enzymatic oxidation.[123] Different reaction conditions were explored to optimize enzyme activity including several reaction buffers, pH values, incubation times and temperatures, but still no hydroxylation was observed.

4.3 DISCUSSION

4.3.1 Isolation of pepticinnamin E

The low yields of pepticinnamin E from the extraction and purification may be due to several factors, including limited availability of tyrosine or phenylalanine precursors in the cell, loss of materials during 5 rounds of purification, and stability of the compound. Pepticinnamin E appears to be unstable and degrade within a few days, because it was no longer detected when stored either at room temperature or −20 °C, and as a dried powder or resuspended in methanol. Finally, several isomers of pepticinnamin E appeared to be present, which have not been reported. Each HPLC fraction containing pepticinnamin E had several peaks, all containing the same 908 mass, and no other reported pepticinnamin molecules have the same *m/z* value.

4.3.2 Expression and characterization of Pcm NRPSs and accessory proteins

There are several possible reasons why neither Pcm2 module is able to activate any of the amino acids tested within our acceptable threshold (9-30% conversion). First, perhaps the presence of the MbtH protein is hindering enzyme activity. It's possible that coexpression with Pcm3 is necessary, which is why Pcm2.1 is much more soluble and abundant than Pcm2.2, but maybe Pcm3 is inducing a conformational change in the Pcm2 modules, rendering them inactive. Secondly, perhaps the modules are not active as designed. The A domains have been cloned and will be purified to test in this assay to overcome this possibility. Lastly, the substrates tested may not be the proper native substrates for either Pcm2 module. Since L-Phe and L-Tyr have the same percent conversion with Pcm2.1, perhaps the native substrate is large and aromatic, but non-canonical. Similarly, since Pcm2.2 only showed very slight activity (0.1%) for D-Tyr, Pcm2.2 may be activating a similar D-amino acid or any other non-canonical substrate. It's very possible that one of these modules requires either *O*-methylated tyrosine, chlorinated tyrosine, or the doubly modified substrate before activation by the NRPS.

4.3.3 Characterization of the tyrosine hydroxylase Pcm11

The lack of Pcm11 activity could be due to several different reaction conditions. One possibility is that Pcm11 acts on T domain-bound tyrosine instead of free tyrosine in solution. Alternatively, the substrate may first be modified by the halogenase Pcm1. Another possibility is the lack of a suitable biopterin cofactor. Mammalian tyrosine hydroxylases require BH4 and 6MPH4 for *in vitro* activity, but perhaps a different biopterin derivative is necessary for Pcm11 activity.

4.4 CONCLUSION AND OUTLOOK

We identified a new bacterial producer of pepticinnamin E and the gene cluster responsible for synthesizing this natural product. We have attempted to reconstitute the activity of the NRPS Pcm2 and the tyrosine hydroxylase Pcm11 under many different conditions. Although neither enzyme has exhibited significant activity under the conditions tested, our efforts have eliminated several reaction conditions and set the stage for future reconstitution.

To obtain pure pepticinnamin E for full NMR analysis and biological studies, we will need to prevent compound degradation by reducing the time necessary for isolation from *A. bacterium* OK006 cultures and purification. The methods optimized for pepticinnamin E will be applied to isolate and characterize the other pepticinnamin variants.

Several experiments are ongoing to characterize the enzymes in the pepticinnamin cluster. Other proteinogenic amino acids will be assayed as substrates for Pcm2 A domains with or without the MbtH protein Pcm3. The putative chlorinase Pcm1 will be purified and assayed for in vitro activity with tyrosine, which will inform the timing of the modifications. The T domains of Pcm2 are currently being expressed and purified for T domain loading assays with tyrosine or derivatives. This T domain-loaded substrate will be tested as the substrate for the hydroxylase Pcm11. The insights we gain will provide a significant step towards elucidating the biosynthesis of pepticinnamin E. Understanding the biosynthesis of this molecule will allow for engineered biosynthesis and development of new molecules with similar scaffolds as potential cancer therapeutics.

4.5 EXPERIMENTAL PROCEDURES

4.5.1 General

All chemicals were purchased from Sigma-Aldrich or Fisher. Primers were ordered from Integrated DNA Technology (Table S4.1). Polymerase, restriction enzymes, and ligase were purchased from New England Biolabs. DNA manipulation kits were purchased from Qiagen and Thermo Scientific. PCR reactions were carried out using an Eppendorf Nexus GSX1 thermocycler with Q5 DNA polymerase. General genetic manipulation of *E. coli* was carried out using standard protocols.[89] LC-MS data was obtained using an Agilent 6520 accurate-mass Q-TOF LC-MS. Compound purification by preparative HPLC was conducted using a Varian Prostar instrument equipped with a PDA detector and compounds are separated on a Phenomenex Luna 10 μm C18 (2) column. Further rounds of purification was carried out using a Shimadzu HPLC equipped with an SPD-20A UV-vis detector and compounds are separated on a Phenomenex Kinetex 5 μm C18 column (100 Å, 150 mm x 4.60 mm). NMR spectra were obtained using a Bruker 600 MHz spectrometer or a Bruker Avance^{III} HD 800 MHz spectrometer unless otherwise noted.

4.5.2 Bacterial strains and culture conditions

For cloning, *E. coli* DH5α and *E. coli* Top10 chemically competent maximum efficiency cells were used (Invitrogen). For heterologous expression, *E. coli* BL21(DE3) (Agilent) or *E. coli* BAP1 electrocompetent cells (made in-house) were transformed with vectors containing expression constructs (Table S4.2). All overnight seed cultures were inoculated with a single colony and grown in 5 mL media in 14-mL Falcon culture tubes. Antibiotics were used in the following concentrations, unless otherwise noted: 10 μg/mL

tetracycline, 25 μg/mL chloramphenicol, 50 μg/mL kanamycin, 50 μg/mL spectinomycin, or 100 μg/mL ampicillin. All *E. coli* cultures for protein expression were grown in LB Media (Miller, granulated, Fisher) at 37 °C to an OD_{600} between 0.4 and 0.8. *E. coli* cultures were induced with 0.5 mM IPTG and then were incubated at 16 °C overnight, while shaking at 225 rpm. *Actinobacteria bacterium* OK006 seed cultures using spore stocks were grown in tryptic soy broth media (17 g/L tryptone, 3 g/L soytone, 2.5 g/L glucose, 5 g/L sodium chloride, and 2.5 g/L potassium phosphate dibasic). Large-scale cultures of OK006 for metabolite extraction were grown in PPM2 media (10 g/L dextrose, 5 g/L L-asparagine, 0.1 g/L magnesium sulfate anhydrous, 0.5 g/L potassium phosphate dibasic, 0.02 g/L iron (II) sulfate heptahydrate, and 5.85 g/L sodium chloride) at 28 °C with shaking at 225 rpm for 5 days and were not induced.

4.5.3 Protein expression and purification

Large-scale cultures for protein purification were grown in 1 L of media containing the appropriate antibiotic and 1−5 mL of overnight culture. At an OD600 of 0.5−0.8, a sample of 0.5 mM IPTG was added to induce protein expression. After 16 hours of protein expression, cell pellets were collected by centrifugation and resuspended in TRIS or phosphate lysis buffer (50 mM TRIS or potassium phosphate dibasic, 200 mM NaCl, 10% glycerol, pH 8). Cells were lysed by sonication at 30% amplitude for 3 minutes (0.5 seconds on and 1.5 seconds off). Cell debris was removed by centrifugation at 17,000 rpm for 40 minutes and the supernatant was filtered using a 0.45 μm filter (Corning). Each Pcm protein, excluding Pcm3, was purified using an AKTA FPLC using a 5 mL HiTrap nickel column (GE) at a flow rate of 3.0 mL/min. Pcm3 was purified using a 5 mL HiTrap Q anion exchange column (GE). Wash buffer consisted of 50 mM TRIS or potassium phoshate, 200 mM NaCl, 25 mM imidazole, and 10%

glycerol (pH 8). Elution buffer was the wash buffer containing 250 mM imidazole. A different buffer system was used to purify Pcm3: wash buffer consisted of 20 mM TRIS (pH 8) and elution buffer consisted of 20 mM TRIS (pH 8) and 1 M NaCl. The wash step lasted for 5 column volumes (CVs). The proteins were eluted using a gradient of 0−100% of elution buffer over 10 CVs, and then held at 100% for 5 CVs. All proteins were further purified using a HiLoad 16/600 Superdex 200 prep grade size exclusion column (GE) on the FPLC using the lysis buffer described above. The column was first equilibrated with lysis buffer at a flow rate of 0.5 mL/min for 1.5 CVs, then protein was loaded onto the column and eluted over 1.5 CVs using the same buffer and flow rate.

4.5.4 Metabolite extraction from A. bacterium OK006

A. bacterium cultures (100 mL) were spun down at 3500 x *g* and 4 °C for 10 minutes. Metabolite extractions from cell pellet were performed using 2 mL methanol per 100 mL culture. Pellets were resuspended in methanol for cell lysis and metabolite extraction, vortexed vigorously and incubated for 2 hours while shaking at room temperature. Cell debris was removed by centrifugation at 3500 x *g* and 4 °C for 20 minutes. The methanol extract was filtered (0.2 µm filter, Corning) into a round bottom flask and concentrated under vacuum. The dried down material was resuspended and transferred to a small glass vial, concentrated under reduced pressure, and stored at −20 °C until ready for MS analysis or purification. Before purification, solvent-solvent extractions were performed. Dry extracts were resuspended in ethyl acetate then water was added. The ethyl acetate layer was collected and concentrated under reduced pressure. Extracts were then resuspended in methanol and hexanes were added. The methanol layer was collected and concentrated under reduced pressure.

4.5.5 Purification of pepticinnamin E

Extracts were resuspended in 5 mL methanol and injected onto the prep HPLC at a flow rate of 15 mL/min with water and acetonitrile (both containing 0.1% trifluoroacetic acid) as mobile phases. The solvent gradient was held at 2% acetonitrile for 5 minutes, ramped up to 58% acetonitrile over 20 minutes and to 95% over another 10 minutes. Fractions containing pepticinnamin E (32–33 min retention time) were concentrated under reduced pressure, resuspended in 3 mL methanol, and injected onto the prep HPLC a second time. A 2−40% acetonitrile solvent gradient over 5 minutes was used, then the gradient was ramped from 40% to 60% over 30 minutes. Fractions containing pepticinnamin E (33.5–34.5 min retention time) were concentrated under reduced pressure, resuspended in 3 mL methanol, and injected onto the prep HPLC a third time. A 2–50% acetonitrile solvent gradient over 5 minutes was used, then the gradient was ramped from 50% to 55% over 10 minutes, then from 55% to 60% over 20 minutes. Fractions containing pepticinnamin E (23–24 min retention time) were concentrated under reduced pressure, resuspended in 200 µL methanol, and injected onto the analytical HPLC using the same mobile phases. At a flow rate of 0.5 mL/min, the gradient ramped from 2% to 45% acetonitrile over 7 minutes, then from 40% to 60% over 30 minutes. The fraction containing pepticinnamin E (22–23 min retention time) was concentrated under reduced pressure, resuspended in 200 µL methanol, and injected onto the analytical HPLC for one final purification. The solvent gradient was ramped from 2% to 45% over 4 minutes, then from 45% to 50% over 5 minutes, then from 50% to 55% over 15 minutes. Pepticinnamin E was detected in the fraction with a retention time of 18.6–19.6 minutes. LC-MS analysis suggested that the final concentrated sample contained ~75% pure pepticinnamin E.

4.5.6 NMR spectroscopy for pepticinnamin E

Pepticinnamin ^{1}H spectra were acquired using a Bruker Avance^{III} HD (600 MHz) instrument equipped with a cryo probe. NMR spectra were processed and baseline corrected using Mestrelabs MNOVA software packages. NMR assignments are shown in Table S4.3.

4.5.7 ATP-$^{32}PP_i$ exchange assays to determine NRPS substrate specificity

Assays contained 50 mM TRIS (pH 9), 2 mM magnesium chloride, 1 mM ATP, 4 mM amino acid being tested, 4 mM "cold" sodium pyrophosphate, 10 mM "hot/cold" sodium pyrophosphate, 4 μM NRPS enzyme, and 5 μM the MbtH protein Pcm3. Assays were diluted with water to a final volume of 100 μL. Reactions were incubated at 37 °C for 45 minutes, then quenched with a solution containing 1.6% activated charcoal, 3.5% perchloric acid, and 0.1 M sodium pyrophosphate. NRPS enzymes tested include full Pcm2, Pcm2 module 1 (Pcm2.1), Pcm2 module 2 (Pcm2.2), and Pcm2.1 A. For each substrate, a no enzyme control and a no MbtH control was included, and a no substrate control was also run. Reactions contained approximately 120,000 cpm $^{32}PP_i$ each.

4.5.8 Pcm11 tyrosine hydroxylase assay

Assays consisted of 100 mM sodium phosphate (pH 6.8), 2 mM DTT, 2.5 mM tyrosine, 0.5 mM biopterin in either 0.005 M HCl or water, 100 μM iron (II) ammonium sulfate, 10 μM Pcm11, and 100 units catalase. Reactions were brought to 100 μL with water. Assays were incubated for 40 minutes at room temperature, quenched with 10% TCA for 10 minutes, neutralized with dilute NaOH, and centrifuged to remove precipitants before LC-MS analysis.

4.5.9 MS analysis

Pcm11 *in vitro* enzyme assays were analyzed by an Agilent 6520 accurate-mass Q-TOF LC/MS using a Phenomenex Kinetex 5 μm C18 column (100 Å, 150 mm x 4.60 mm). Mobile phases were water and acetonitrile, each containing 0.1% formic acid. The gradient was held at 2% acetonitrile for 5 minutes before ramping up to 45% acetonitrile over 4 minutes and then 95% acetonitrile over 4 minutes at a flow rate of 0.4 mL/min.

APPENDIX: SUPPLEMENTAL TABLES AND FIGURES

Supplemental Tables

Table S2.1. Gene names and protein identifiers of the biosynthetic operons responsible for production of phevamine and the toxins in Figure 2.4.

Toxins	Gene	Protein Identifier
Tabtoxin	tabP	AAL99263
	tabD	AAM13981
	tabB	AAB41803
	tabA	AAB41802
	tblA	AAB25380
	tabC	AAM75344
	tblS	AAM77668
	tblC	AAO63153
	tblD	AAU95210
	tblE	AAP13070
	tblF	AAP13071
	tblR	AAP13072
Phaseolotoxin	phtT	AAZ37741
	phtS	AAZ35586
	phtO	AAZ36603
	amtA	AAZ35074
	phtN	AAZ36976
	phtM	AAZ37830
	phtL	BAF32890
	desI	AAZ37112
	phtD	AAZ36338
	phtB	AAZ37805
	phtA	AAZ35592
	argK	AAZ35351
Syringolin	sylA	ELS43795
	sylB	ELS43794
	sylC	ELS43793
	sylD	ELS43792
	sylE	ELS43791
Phevamine	hsvA	NP_790712.1
	hsVB	NP_790713.1
	hsvC	NP_790714.1

Mangotoxin	mboA	
	mboB	
	mboC	
	mboD	
	mboE	
Coronatine	hhd	
	ocs	
	acsl	
	Cfa6	
	Cfa7	
	ccc	
	mcah	
	LuxRNew	
	CmaA	
	CmaT	

Table S2.2. **Primers used for cloning of the *hsv* and *mbo* operons, cloning of individual *hsv* and *mbo* genes in *E. coli*, and generation of *hsv* knockouts.** Nucleotides added for cloning are shown in bold. Restriction sites are underlined.

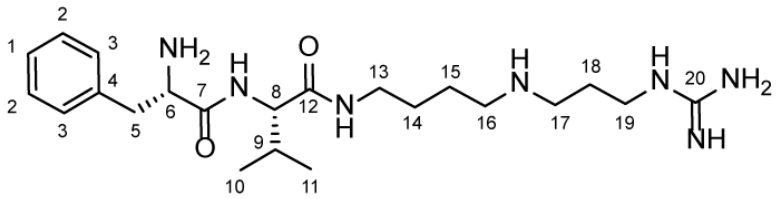
Name	Sequences (5' to 3')
hsv_pLIC-His_sense	TACTTCCAATCCAATGCGATGTCATATCAAAAAGCAGAACCCGCTTAC
hsv_pLIC-His_anti	TTATCCACTTCCAATGCGCTATCAACCACGCGCAGCTTCTGCATGC
hsv_pBBR5_sense	GACAACAGACTAGTATGTCATATCAAAAAGCAGAACCCG (SpeI site underlined)
hsv_pBBR5_anti	GACAACAGAAGCTT TCAACCACGCGCAGCTTC (HindIII site underlined)
hsvA_pRSF_sense	GACAACAGGGATCCGATGTCATATCAAAAAGCAGAAC (BamHI site underlined)
hsvA_pRSF_anti	GACAACAGAAGCTTTCAAATATAACGGTGCAGTCCT (HindIII site underlined)
hsvB_pACYC_sense	GACAACAGGAGCTCGATGCGCCCTACAAAAAAAATACTG (SacI site underlined)
hsvB_pACYC_anti	GACAACAGAAGCTTTCACTCCTCAAACGGAGGG (HindIII site underlined)
hsvC_pCDF_sense	GACAACAGGGATCCGATGGACAAAAATAAGCCAAACACTATTCTG A (BamHI site underlined)
hsvC_pCDF_R	GACAACAGAAGCTTTCAACCACGCGCAGCTTCTG (HindIII site underlined)
MT1797 (hsv)	GTCCCGGATGAAGTAATCGGATCC
MT1881 (hsv)	GGGTACTGTCTAACGCCGCTATATTGACTGAAGGGGCG

MT1880 (hsv)	ATATAGCGGCGTTAGACAGTACCCGTCCGGGCATCCC
MT1877 (hsv)	CACCGTCTGAGGTTCTGATAGGACGGGG
MT1878 (hsv)	CTGCGCGACGAAATGCTCGCGGCGG
MT1967 (hsv)	GGGGTTCGTCCGGCCTTTCCGC
MT1968 (hsv)	GTGCTGATTGCGAGCACGGTCG
MT1969 (hsv)	GCTTCAGACCTTCCTCAAGGCG
MT1970 (hsv)	GTCTAATGGCTGACCGCCACTG
MT1971 (hsv)	AGAACCCACGTAGTGCCGGCTC
MT1972 (hsv)	ATGGACGCCATAGGGTGTGTAG
mboC_pCDF_sense	GACAACAGGGATCCGATGTCCATCGTTGTCATTGATC CT (BamHI site underlined)
mboC_pCDF_anti	GACAACAGAAGCTT TCATACGATTCCAACCTCTGTGTG (HindIII site underlined)
mboD_pACYC_sense	GACAACAGGGATCCGATGAAGCATGTCCTGATCATC AACCGC (BamHI site underlined)
mboD_pACYC_anti	GACAACAGAAGCTTTCAGGTTGCGCGAATAGACGTC TG (HindIII site underlined)
mboE_pRSF_sense	GACAACAGGGATCCGATGAAAGTGCTCCTGCAGCTT (BamHI site underlined)
mboE_pRSF_anti	GACAACAGAAGCTT TTATAGGAGTTGAGTGGCGCAATG (HindIII site underlined)
mboA-F_pET52LIC_sense	CAGGGACCCGGTTCTAGAATGATATCCACCGCTGAG ATG CTA (XbaI site underlined)
mboA-F_pET52LIC_anti	GGCACCAGAGCGTTGCGGCCGCTCAGGCCTCGCTGG GAGT (NotI site underlined)
mboDEF_pRSF_sense	GACAACAGGGATCCGATGAAGCATGTCCTGATCATC AACCGC (BamHI site underlined)
mboDEF_pRSF_anti	GACAACAGAAGCTTTCAGGCCTCGCTGGGAGTCTT (HindIII site underlined)
mboB_pRSF_sense	GACAACAGGGATCCGATGGCAATCTACAGTGTTCGT ATT (BamHI site underlined)
mboB_pRSF_anti	GACAACAGAAGCTTCTATCCAAAACTTTCAAGTTCAC ACAA (HindIII site underlined)

Table S2.3. List of plasmids and strains used in the *hsv* and *mbo* studies.

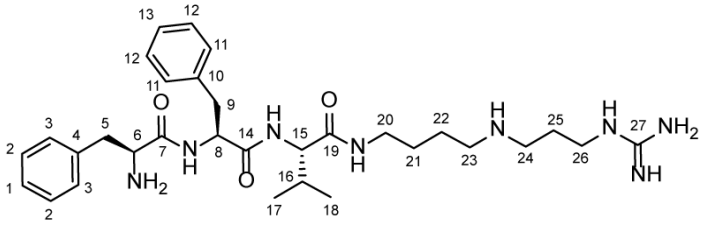
Insert	Plasmid	Host strain	Source
hsvA	pRSF-duet (Kan)	*E. coli* BL21	This study
hsvB	pACYC-duet (Cam)	*E. coli* BL21	This study
hsvC	pCDF-duet (Sm)	*E. coli* BL21	This study
hsvA-C	pLIC-His (Amp)	*E. coli* BL21	This study
		*P. fluorescens Pf*0-1	63
hrp/hrc		*EthAn Tn5 transposant of P. fluorescens Pf*0-1 (Tet)	64
hsvA-C	pBBR5 (Gent)	*P. fluorescens Pf*0-1	This study
		P. syringae pv. *tomato* DC3000	13
hrpL	pBAD (Tet)	*P. syringae* pv. *tomato* DC3000	13
*Tn*5::cfa6 (DC3118)		*P. syringae* pv. *tomato* DC3000	124-125
mboB	pRSF-duet (Kan)	*E. coli* BL21	This study
mboC	pCDF-duet (Sm)	*E. coli* BL21	This study
mboD	pACYC-duet (Cam)	*E. coli* BL21	This study
mboE	pRSF-duet (Kan)	*E. coli* BL21	This study
mboA-F	pET52 3C/LIC (Amp)	*E. coli* BL21	This study
mboDEF	pRSF-duet (Kan)	*E. coli* BL21	This study
		P. syringae pv. s*yringae* UMAF0158	126

Table S2.4. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for phevamine A in D_2O. Chemical shifts were referenced to δ($C\underline{H}_3CO_2H$) = 2.08 ppm and δ($^{13}\underline{C}H_3CO_2H$) = 21.03. ^{13}C chemical shifts were determined via HMBC and HSQC experiments. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the indicated proton(s) to the indicated ^{13}C atom.



Position	δ_c	Proton	δH(*J*$_{HH}$[Hz])	HMBC
1	128.49	1-H	7.41($J_{1,2}$ = 6.9)	3
2	129.61	2-H	7.42($J_{2,1}$ = 6.9) ($J_{2,3}$ = 7.1)	2,4
3	129.86	3-H	7.26($J_{3,2}$ = 7.1)	1,3,5
4	134.07			
5	37.38	5-Ha	3.19($J_{5a,5b}$ = 14.4) ($J_{5a,6}$ = 7.5)	3,4,6,7
		5-Hb	3.23($J_{5b,5a}$ = 14.4) ($J_{5b,6}$ = 6.6)	3,4,6,7
6	54.65	6-H	4.33($J_{6,5a}$ = 7.5) ($J_{6,5b}$ = 6.6)	4,7
7	169.28			
8	60.45	8-H	3.98($J_{8,9}$ = 8.5)	7,9,10,11,12
9	30.71	9-H	1.95($J_{9,8}$ = 8.5) ($J_{9,10}$ = 6.8) ($J_{9,11}$ = 6.8)	8,10,11
10	18.49	10-3H	0.93($J_{10,9}$ = 6.8)	8,9,11
11	18.75	11-3H	0.90($J_{11,9}$ = 6.8)	8,9,10
12	172.58			
13	39.18	13-Ha	3.15($J_{13a,13b}$ = 13.3) ($J_{13a,14}$ = 7.0)	12,14,15
		13-Hb	3.29($J_{13b,13a}$ = 13.3) ($J_{13b,14}$ = 7.0)	12,14,15
14	26.02	14-2H	1.61($J_{14,13a}$ = 7.0) ($J_{14,13b}$ = 7.0) ($J_{14,15}$ = 7.6)	13,15,16
15	23.65	15-2H	1.73($J_{15,14}$ = 7.6) ($J_{15,16}$ = 7.6)	13,14,16
16	47.74	16-2H	3.10($J_{16,15}$ = 7.6)	14,15,17
17	45.29	17-2H	3.12($J_{17,18}$ = 8.0)	16,18,19
18	25.47	18-2H	2.00($J_{18,17}$ = 8.0) ($J_{18,19}$ = 6.7)	17,19
19	38.71	19-2H	3.31($J_{19,18}$ = 6.7)	17,18,20
20	157.38			

Table S2.5. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for phevamine B in D_2O. Chemical shifts were referenced to δ($C\underline{H}_3OH$) = 3.34 ppm and δ($^{13}\underline{C}H_3OH$) = 49.50. ^{13}C chemical shifts were determined via HMBC and HSQC experiments. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the indicated proton(s) to the indicated ^{13}C atom.



Position	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1	128.62	1-H	7.35($J_{1,2}$ = 7.0)	3
2	129.78	2-2H	7.38($J_{2,1}$ = 7.0) ($J_{2,3}$ = 7.1)	2,4
3	129.99	3-2H	7.23($J_{3,2}$ = 7.1)	1,3,5
4	134.05			
5	37.37	5-2H	3.18($J_{5,6}$ = 7.0)	3,6,7
6	54.66	6-H	4.25($J_{6,5}$ = 7.0)	4,7
7	169.12			
8	55.56	8-H	4.64($J_{8,9a}$ = 8.7) ($J_{8,9b}$ = 7.0)	7,9,13
9	37.93	9-Ha	2.99($J_{9a,9b}$ = 13.2) ($J_{9a,8}$ = 8.7)	10,11
		9-Hb	3.04($J_{9b,9a}$ = 13.2) ($J_{9b,8}$ = 7.0)	10,11
10	136.40			
11	129.78	11-2H	7.20($J_{11,12}$ = 7.5)	9,11,13
12	129.35	12-2H	7.34($J_{12,11}$ = 7.5) ($J_{12,13}$ = 7.0)	10,12
13	127.84	13-H	7.30($J_{13,12}$ = 7.0)	11
14	172.18			
15	60.39	15-H	3.85($J_{15,16}$ = 8.4)	14,16,18,19
16	30.81	16-H	1.91($J_{16,15}$ = 8.4) ($J_{16,17}$ = 6.8) ($J_{16,18}$ = 6.8)	18
17	18.87	17-3H	0.87($J_{17,16}$ = 6.8)	15,16,18
18	18.72	18-3H	0.91($J_{18,16}$ = 6.8)	15,16,17
19	172.85			
20	39.20	20-Ha	3.08($J_{20a,20b}$ = 13.5) ($J_{20a,21}$ = 7.0)	19,21,22
		20-Hb	3.18($J_{20b,20a}$ = 13.5) ($J_{20b,21}$ = 7.0)	19,21,22
21	26.07	21-2H	1.54($J_{21,20a}$ = 7.0) ($J_{21,20b}$ = 7.0) ($J_{21,22}$ = 7.0)	20,23
22	23.69	22-2H	1.67($J_{22,21}$ = 7.0) ($J_{22,23}$ = 7.8)	20,23
23	47.80	23-2H	3.05($J_{23,22}$ = 7.8)	21,22,24
24	45.35	24-2H	3.09($J_{24,25}$ = 7.8)	23,25,26

25	25.56	25-2H	1.98($J_{25,24}$ = 7.8) ($J_{25,26}$ = 7.0)	24,26
26	38.80	26-2H	3.28($J_{26,25}$ = 7.0)	24,25,27
27	157.47			

Table S3.1. Primers used for cloning of the *kwf* cluster and individual *kwf* genes in *E. coli.*

Nucleotides added for cloning are shown in bold. Restriction sites are underlined.

Name	Sequences (5' to 3')
orfBAT_pET30a_sense	GACAACAGCATATGGGTATCGGTTCAAAAACCA ATTCAAAT (NdeI site underlined)
orfBAT_pET30a_anti	GACAACAGGCGGCCGCTCAGGCTGTTAGGTACT GCTCTGATAC (NotI site underlined)
T7Pro_pETcoco_sense	TGCTTGGAGTTGAGGTGGTGTGATAATACGACT CACTATAGGGGAAT
orfBAT_pETcoco_anti	CTGGCTGGCTTGCGGCCGCATCAGGCTGTTAGG TACTGCTC
orfCDE_pETcoco_sense	GCCACGTGCGTACGGTTTCGATTGAGTACTGAA AAAGGGTCC
orfCDE_pETcoco_anti	CCAACTTTTGGAGTTTTCCATTCATTGTTCATCC GCCAGTTC
orfFGHIJK_pETcoco_sense	GAACTGGCGGATGAACAATGAATGGAAAACTCC AAAAGTTGG
orfFGHIJK_pETcoco_anti	CCCCTATAGTGAGTCGTATTATCACACCACCTCA ACTCCAAG
orfB_pLICMBP_sense	TACTTCCAATCCAATGCGATGGGTATCGGTTCAA AAACCAATTCAAAT
orfB_pLICMBP_anti	TTATCCACTTCCAATGCGCTATCAACGAGCACCG CCTGC
orfE_pLICMBP_sense	TACTTCCAATCCAATGCGATGAACAGCACCTTTG AAATAACTAATTATGTGCTC
orfE_pLICMBP_anti	TTATCCACTTCCAATGCGCTATCATTGTTCATCC GCCAGTTCTGC
orfD-tagless_pRSF_sense	GATACACCATGGGCGATAGCAAAAGCATGTATG ACGTTGTCATC (NcoI site underlined)
orfD-tagless_pRSF_anti	GATACAAAGCTTTTACAGCCCCCCGACCGACAG (HindIII site underlined)
orfF_pET30a_sense	GACAACAGGGTACCATGGAAAACTCCAAAAGTT GGCGT (KpnI site underlined)
orfF_pET30a_anti	GACAACAGAAGCTTTCATTTGATCGGATAGTTA TCAACCCAG (HindIII site underlined)
orfT_pRSF_sense	GATACAGGATCCGATGACCATCGAGATCCAAAC TCTGGAAGAC (BamHI site underlined)
orfT_pRSF_anti	GATACAGCGGCCGCTCAGGCTGTTAGGTACTGC TCTGATACTGC (NotI site underlined)

orfE1-ATC_pRSF_sense	GACAACAGGGTACCATGTTGAAGCTGGCGAGGGGA (KpnI site underlined)
orfE1-ATC_pRSF_anti	GACAACAGAAGCTTTCAAGCGGCTTCCGACAGAGCC (HindIII site underlined)
orfE2-A_pRSF_sense	GATACAGGATCCGATGCTTCTGGTGTCGCCACCGGTT (BamHI site underlined)
orfE2-A_pRSF_anti	GATACAAAGCTTTCATTTCAGCGCCGCTTCGATTTC (HindIII site underlined)
orfB-A_pRSF_sense	GATACAGATATCGATGACGCTAGACGAGCAGGTCTGG (EcoRV site underlined)
orfB-A_pRSF_anti	GATACAGGTACCTCACACTGCCTCGATTTCGCCTGG (KpnI site underlined)
orfB-T_pRSF_sense	GATACAGGATCCGATGCTGCCTTCGGTAATAACCGT (BamHI site underlined)
orfB-T_pRSF_anti	GATACAAAGCTTTCAATCAATGCGCGTGAAGTCGAT (HindIII site underlined)

Table S3.2. List of plasmids and strains used in the *kwf* study.

Insert	Plasmid	Host strain	Source
orfCDEFGHIJKBAT	pETcoco-1 (Cam)	*E. coli* BAP1	This study
		P. syringae pv. *actinidiae* NZV-13	[72]
ΔB		*P. syringae* pv. *actinidiae* NZV-13	This study
ΔE		*P. syringae* pv. *actinidiae* NZV-13	This study
ΔB ΔE		*P. syringae* pv. *actinidiae* NZV-13	This study
orfB A domain	pRSF-duet (Kan)	*E. coli* BL21	This study
orfB T domain	pRSF-duet (Kan)	*E. coli* BL21	This study
*orfE*1 A-T-C	pRSF-duet (Kan)	*E. coli* BL21	This study
*orfE*2 A domain	pRSF-duet (Kan)	*E. coli* BL21	This study
orfT domain	pRSF-duet (Kan)	*E. coli* BL21	This study
orfA	pLIC-His (Amp)	*E. coli* BL21	This study
orfB	pLIC-MBP (Amp)	*E. coli* Bap1	This study
orfD	pRSF-duet (Kan)	*E. coli* BL21	This study
orfE	pLIC-MBP (Amp)	*E. coli* Bap1	This study
orfF	pET30a (Kan)	*E. coli* BL21	This study
orfG	pLIC-His (Amp)	*E. coli* BL21	This study
orfH	pLIC-His (Amp)	*E. coli* BL21	This study
orfI	pLIC-His (Amp)	*E. coli* BL21	This study
orfJ	pLIC-His (Amp)	*E. coli* BL21	This study
orfBAT	pET30a (Kan)	*E. coli* Top10	This study

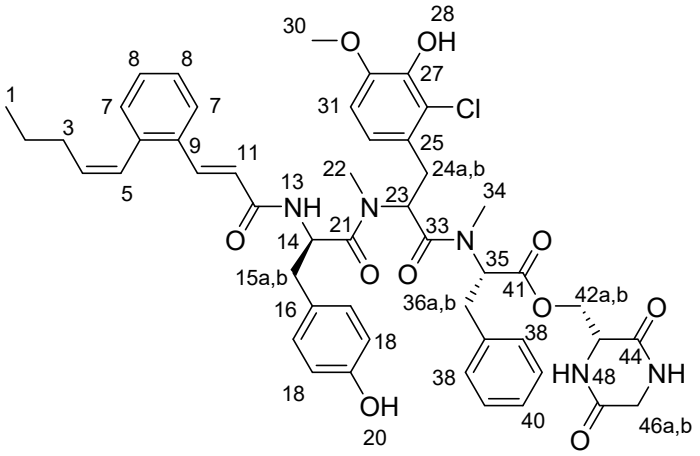
Table S4.1. Primers used for cloning of individual *pcm* genes in *E. coli.* Nucleotides added for cloning are shown in bold. Restriction sites are underlined.

Name	Sequences (5' to 3')
Pcm2.2_pLICMBP_sense	TACTTCCAATCCAATGCGATGGTGCTGCACGAG TGGAAC
Pcm2.2_pLICMBP_anti	TTATCCACTTCCAATGCGCTAGTCACCCTTCTGA CGGGCAC
Pcm2.1A_pRSF_sense	GATACAGGATCCGATGGCGGACGACCCGGCC (BamHI site underlined)
Pcm2.1A_pRSF_anti	GATACAAAGCTTTCAGACGACCACGGCGACCTC TCG (HindIII site underlined)
Pcm2.1T_pRSF_sense	GATACAGGATCCGATGGTGCCCGCGGCGGTC (BamHI site underlined)
Pcm2.1T_pRSF_anti	GATACAAAGCTTTCACACGGGAGGGCGCGGCGA (HindIII site underlined)
Pcm2.2A-MT_pRSF_sense	GATACACATATGGTGGCGAATCCCGAGG (NdeI site underlined)
Pcm2.2A-MT_pRSF_anti	GATACAGATATCTCACTCGGTGTCGATCCGCTCG (EcoRV site underlined)
Pcm2.2T_pRSF_sense	GATACAGGATCCGATGGTCCCGACGGCCCTG (BamHI site underlined)
Pcm2.2T_pRSF_anti	GATACAAAGCTTTCAGTCACCCTTCTGACGGGC (HindIII site underlined)

Table S4.2. List of plasmids and strains used in the *pcm* study.

Insert	Plasmid	Host strain	Source
		A. bacterium OK006	[127]
pcm2.2	pLIC-MBP (Kan)	*E. coli* Bap1	This study
pcm2.1 A	pRSF-duet (Kan)	*E. coli* BL21	This study
pcm2.1 T	pRSF-duet (Kan)	*E. coli* BL21	This study
pcm2.2 A-MT	pRSF-duet (Kan)	*E. coli* BL21	This study
pcm2.2 T	pRSF-duet (Kan)	*E. coli* BL21	This study
pcm11	pET30a (Kan)	*E. coli* BL21	This study
pcm3	pCDF-duet (Spec)	*E. coli* BL21	This study

Table S4.3. ^{1}H (600 MHz) NMR spectroscopic data for pepticinnamin E in MeOD. Protons are assigned based on chemical shift and peak splitting. Protons not observed in MeOD are labeled N/A. Protons that cannot be individually assigned based on ^{1}H NMR alone are grouped and the range of their chemical shifts is listed.



Pepticinnamin E

Position	Proton	δH	Multiplicity
1	1-3H	0.89	t
2	2-2H	1.42	h
3	3-2H	2.04	q
4	4-H	5.58	q
6			
9			
12			
13	13-H	N/A	
16			
19			
20	20-H	N/A	
21			
22	22-3H	3.59	s
25			
26			
27			
28	28-H	N/A	
29			
30	30-3H	3.66	s
33			
34	34-3H	3.21	s
37			

41 44 45 47 48	45-H 48-H	N/A N/A
15, 24, 36	15-2H, 24-2H, 36-2H	3.13-3.67
14, 23, 35, 42, 43, 46	14-H, 23-H, 35-H, 42-2H, 43-1H, 46-2H	4.40-5.08
5, 7, 8, 10, 11, 17, 18, 31, 32, 38, 39, 40	5-H, 7-2H, 8-2H, 10-H, 11-H, 17-2H, 18-2H, 31-H, 32-H, 38-2H, 39-2H, 40-H	6.45-7.80

Supplemental Figures

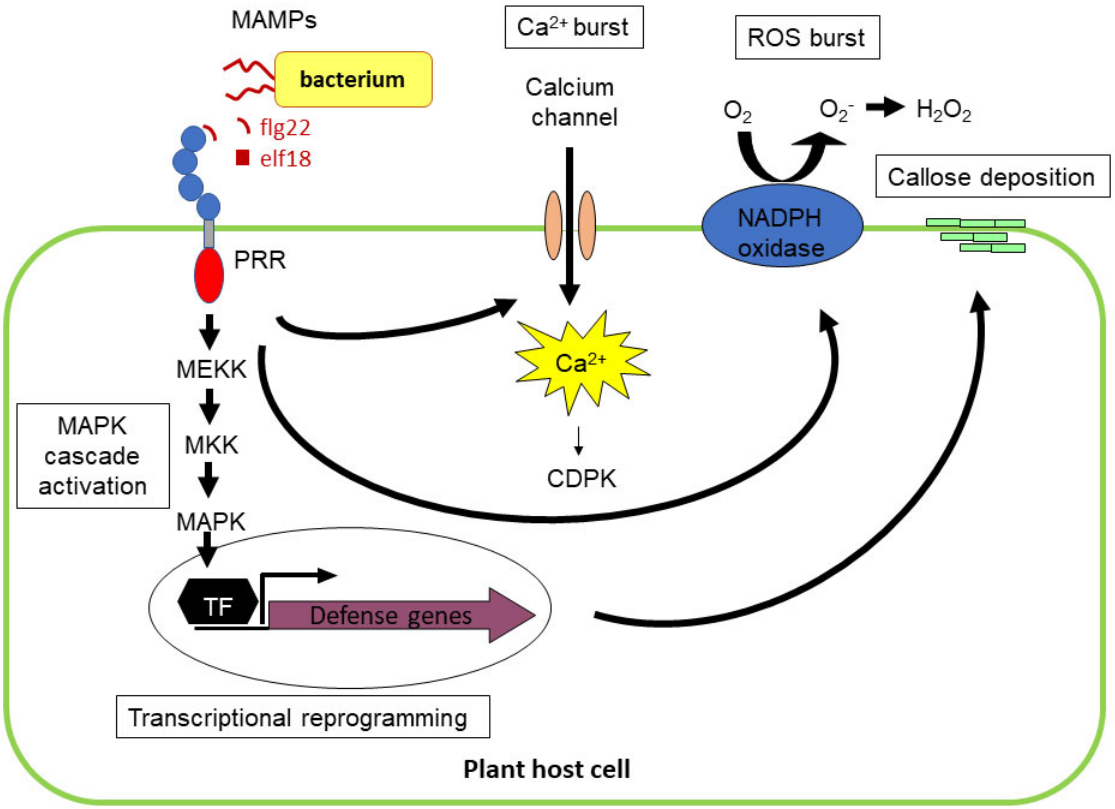


Figure S2.1. Plant immune response schematic. Diagram showing several MAMP-induced immune responses in a plant host cell. Upon elicitation by a MAMP, a plant cell may induce MAPK cascade activation, transcriptional reprogramming, calcium burst, ROS burst, synthesis of various secondary metabolites, and callose deposition. Collectively, these responses are sufficient to suppress pathogen proliferation. MAMP: microbe-associated molecular pattern, PRR: pathogen recognition receptor, TF: transcription factor, MEKK: mitogen-activated protein (MAP) kinase kinase kinase, MKK: MAP kinase kinase, MAPK: MAP kinase, CDPK: calcium-dependent protein kinase, ROS: reactive oxygen species.

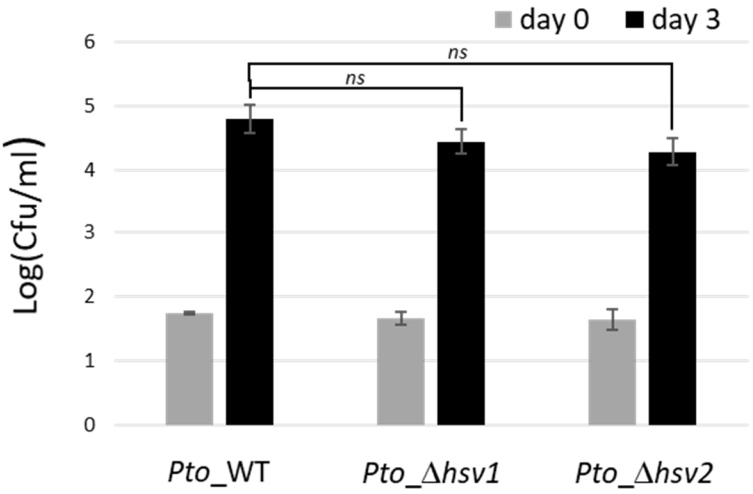


Figure S2.2. *Pto*_Δ*hsv* mutants do not display significantly reduced virulence on *A. thaliana*. Bacteria were counted after recovery from *Pto*-wildtype and *Pto*_Δ*hsv* mutants on *A. thaliana* seedlings. The experiment was repeated twice with similar results. The mutants in both experiments displayed slightly reduced growth but this was not statistically significant (Student's *t*-tests). ns: non significant. Error bars represent ± standard error.

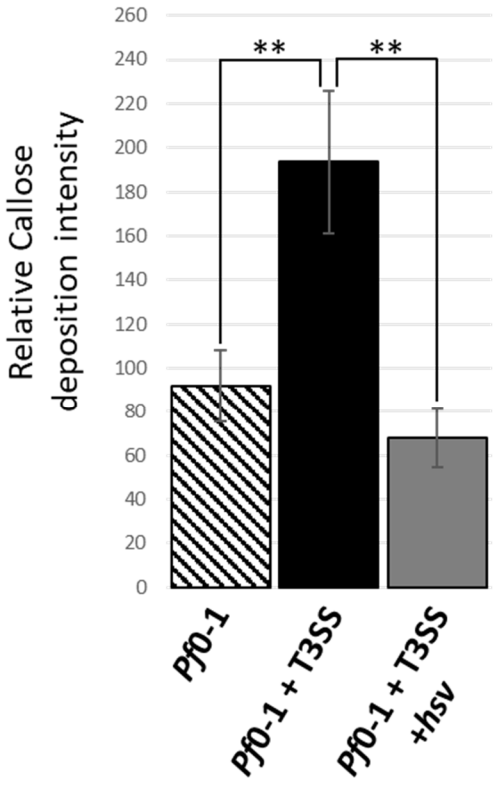


Figure S2.3. The *hsv* operon suppresses callose deposition in *A. thaliana* induced by components encoded by the T3SS locus. Leaves of 4-week-old *A. thaliana* were infiltrated with the *P. fluorescens* strains noted at bottom and stained with aniline blue, ~20 h post infiltration. **, p-value ≤ 0.01; t-test. Error bars represent standard errors. This experiment was repeated 3 times with similar results.

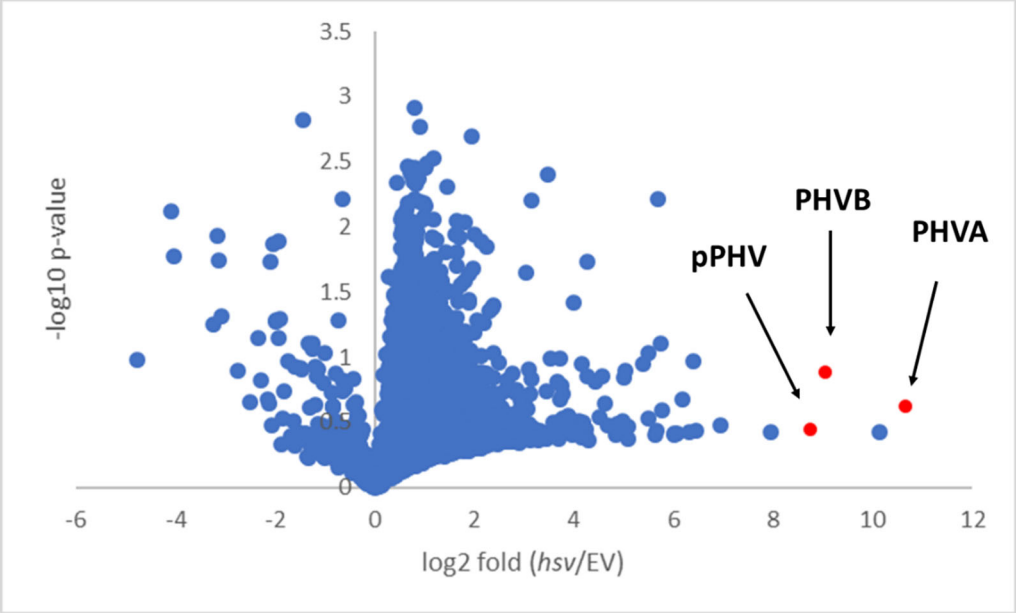


Figure S2.4. Identification of phevamines by comparative metabolomics. The LC-MS metabolic profiles are compared between the extract of *E. coli* overexpressing *hsv* and *E. coli* control containing the empty vector. The fold change ($\log_2$) for each metabolite is plotted against the *p*-value ($\log_{10}$) for that metabolite. Data points for pre-phevamine (pPHV), phevamine A (PHVA), and phevamine B (PHVB) are highlighted in red. The average fold change is very large for pre-phevamine (430 fold increase), phevamine A (1600 fold increase), and phevamine B (540 fold increase) comparing *hsv* overexpression and the empty vector control. When the fold change for a metabolite is large, the level in each sample tends to vary more. The relatively large *p*-values for these metabolites may be due to variable protein levels upon overexpression.

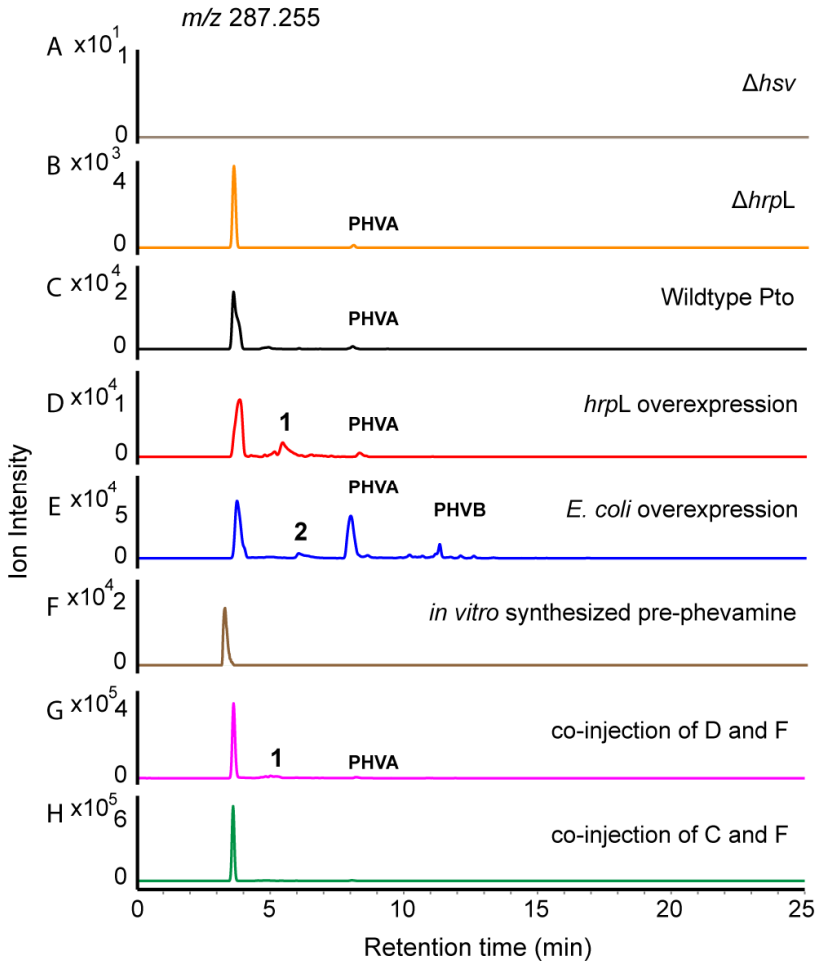


Figure S2.5. **Production of pre-phevamine by different bacterial strains.** Extracted ion chromatograms are shown for pre-phevamine (*m/z*, 287.255) detected in the extracts of each bacterial strain. The retention time of pre-phevamine is approximately 3.8 minutes. Pre-phevamine is not produced by *Pto*Δ*hsv* (*hsv* deletion mutant in *Pto*) (gray). Pre-phevamine is produced by *Pto*Δ*hrpL* (*hrpL* deletion mutant in *Pto*) (orange), wildtype *Pto* (black), *Pto* pBAD:*hrpL* overexpression (red), and *E. coli* overexpressing the *hsv* operon (blue). Pre-phevamine is also produced in the *in vitro* enzymatic reaction and was purified by HPLC (brown). The *in vitro* purified standard was co-injected with the *hrpL* overexpression extract (pink) and the wildtype *Pto* extract (green), both showing a single peak for pre-phevamine. In (*D*), (*E*), and (*G*), the additional pre-phevamine peaks result from in-source fragmentation of phevamine A or phevamine B as labeled. Lowering the fragmentor voltage reduced these in-source fragmentation peaks. Two metabolites from extractions, 1 and 2, are structurally unrelated to pre-phevamine based on LC/MS/MS analysis.

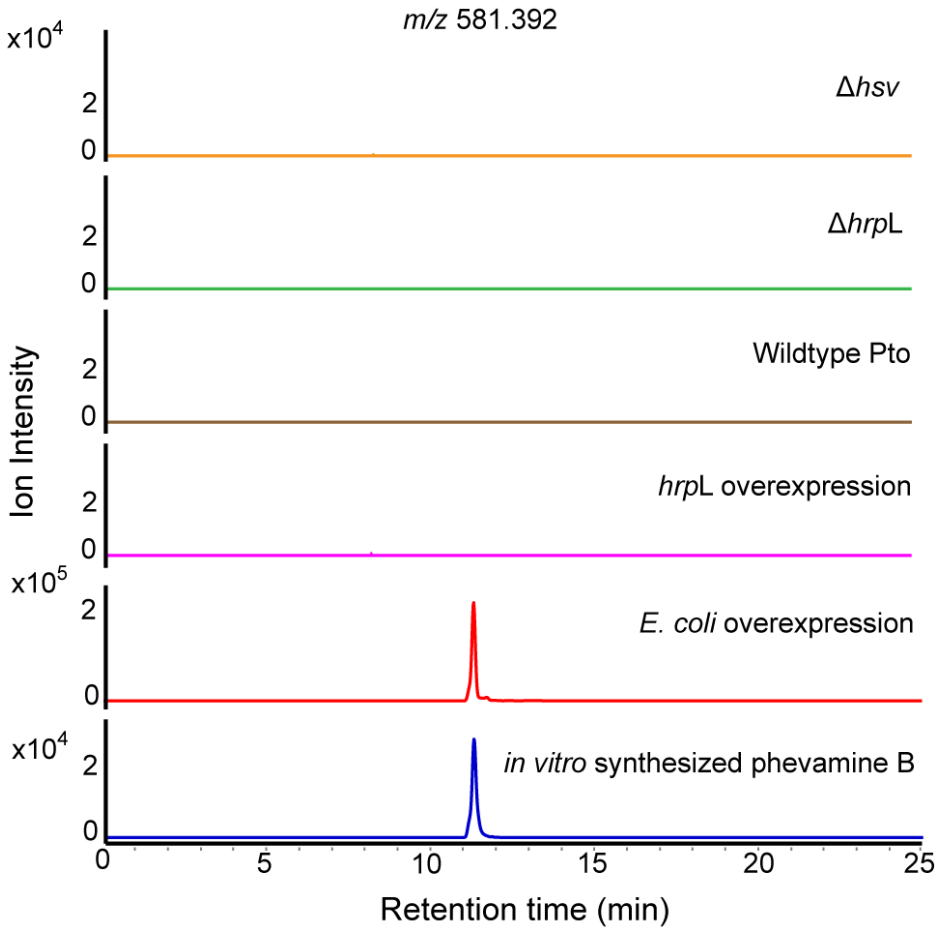


Figure S2.6. Production of phevamine B by different bacterial strains. Extracted ion chromatograms are shown for phevamine B (*m/z*, 581.392). Phevamine B is not produced by *Pto*Δ*hsv* (*hsv* deletion mutant in *Pto*) (orange), *Pto*Δ*hrpL* (*hrpL* deletion mutant in *Pto*) (green), wildtype *Pto* (brown), or *Pto* pBAD:*hrpL* overexpression (pink). Phevamine B is only detected in *E. coli* overexpressing *hsv* (red) and *in vitro* enzymatic reaction (blue).

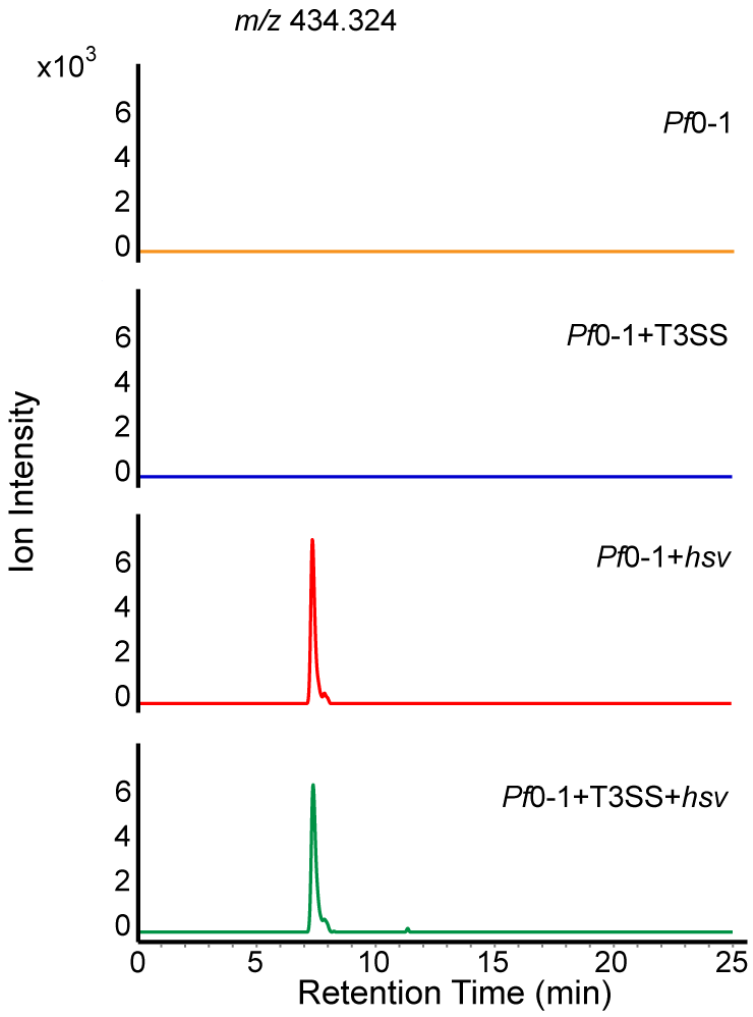


Figure S2.7. Phevamine A production in *P. fluorescens Pf*0-1 is independent of T3SS. Extracted ion chromatograms of phevamine A (*m/z* 434.324) are shown. Phevamine A is detected in both *Pf*0-1 strains expressing *hsv, Pf*0-1+T3SS+*hsv* (green) and *Pf*0-1+*hsv* (red), indicating that the T3SS is not required for the secretion of phevamine A. Phevamine A is not detected in the culture extracts of control strains without *hsv*, *Pf*0-1+T3SS (blue) or *Pf*0-1 (orange).

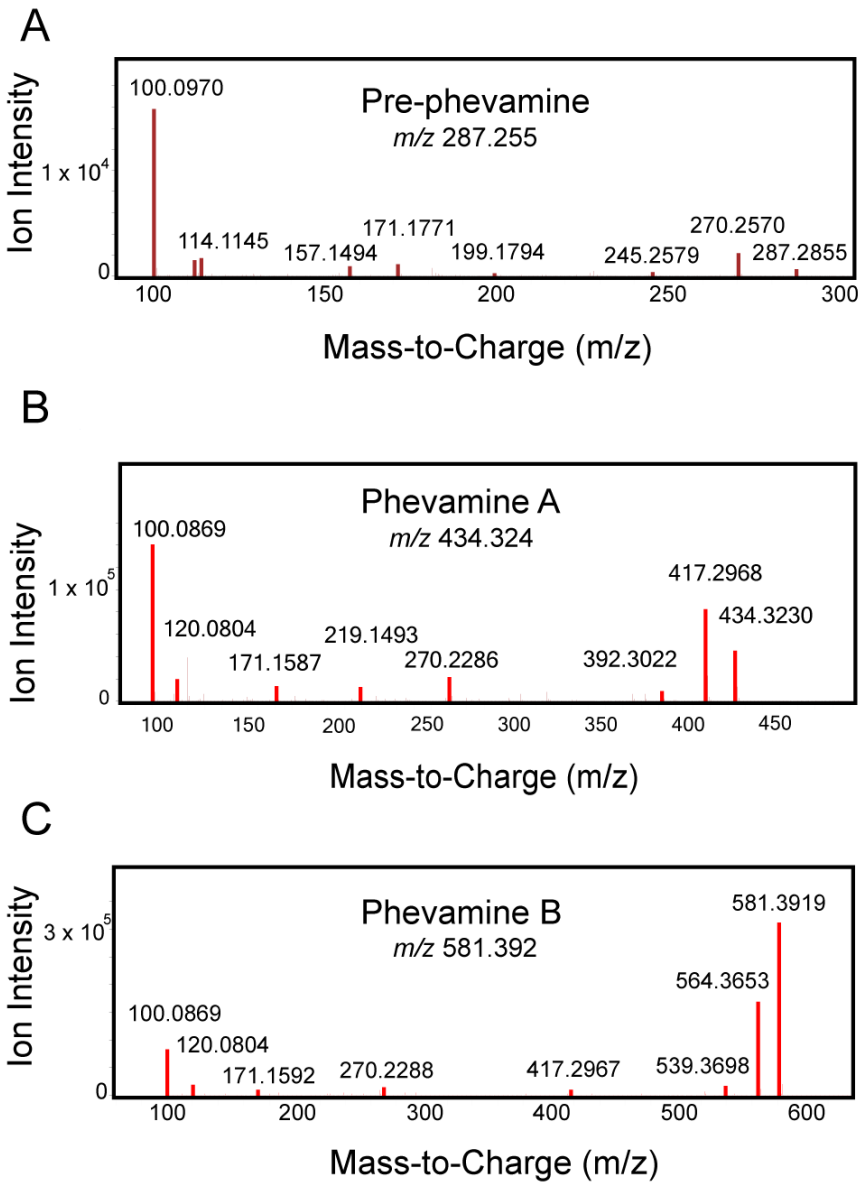


Figure S2.8. Structural analyses of pre-phevamine and phevamines by tandem mass spectrometry (LC-MS/MS). *(A)* MS/MS spectrum of pre-phevamine (*m/z* 287.255). *(B)* MS/MS spectrum of phevamine A (*m/z* 434.324). (*C)* MS/MS spectrum of phevamine B (*m/z* 581.392).

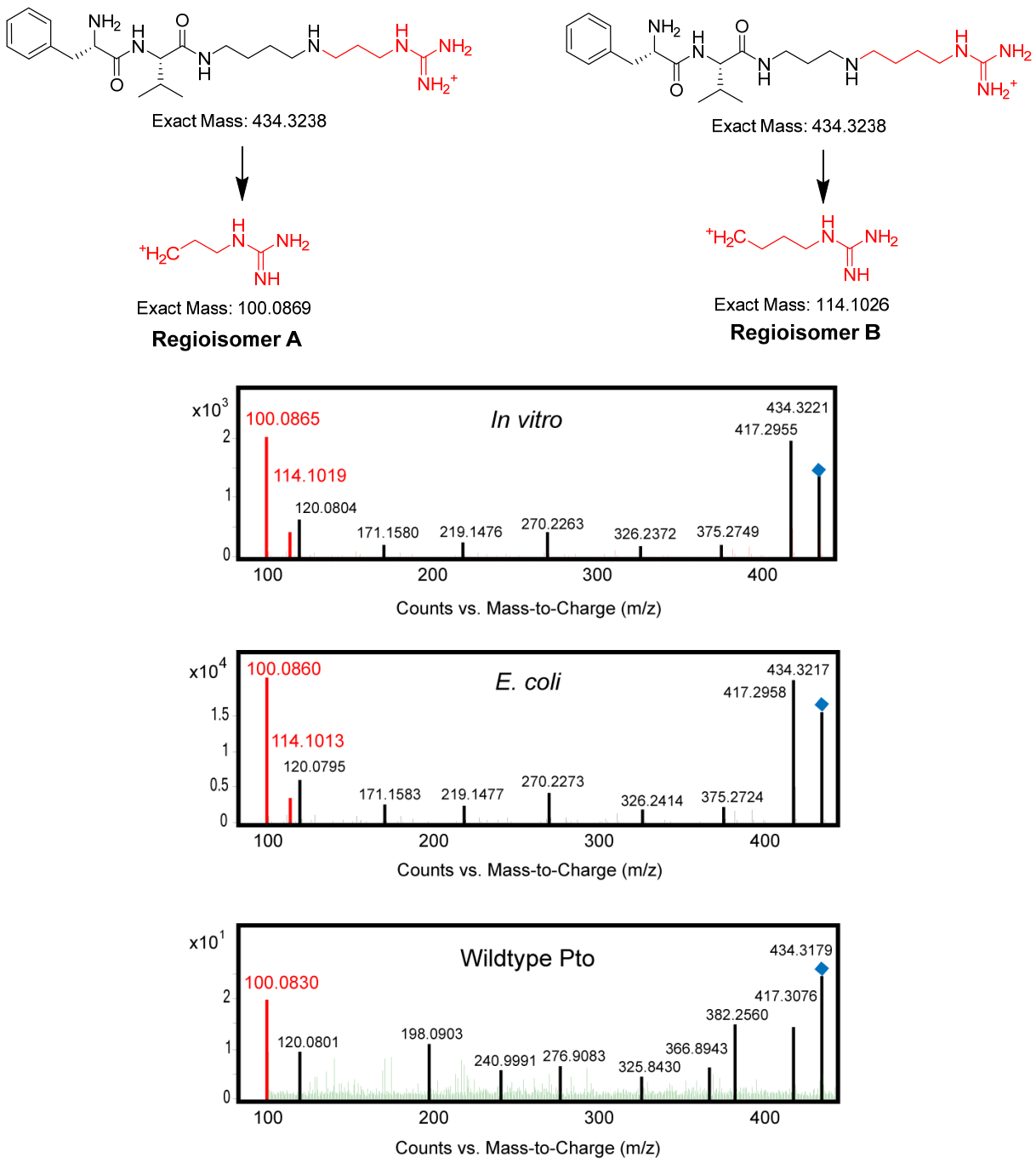


Figure S2.9. Identification of the regioisomers of phevamine A by LC-MS/MS. A 100 *m/z* fragment is unique to the amidinotransfer product on the propylamine end of spermidine (regioisomer A), and a 114 *m/z* fragment is unique to the amidinotransfer product on the butylamine end of spermidine (regioisomer B). Both isomers are detected from the *in vitro* reaction and *E. coli* overexpression. In culture extracts of wildtype *Pto*, only regioisomer A was detected.

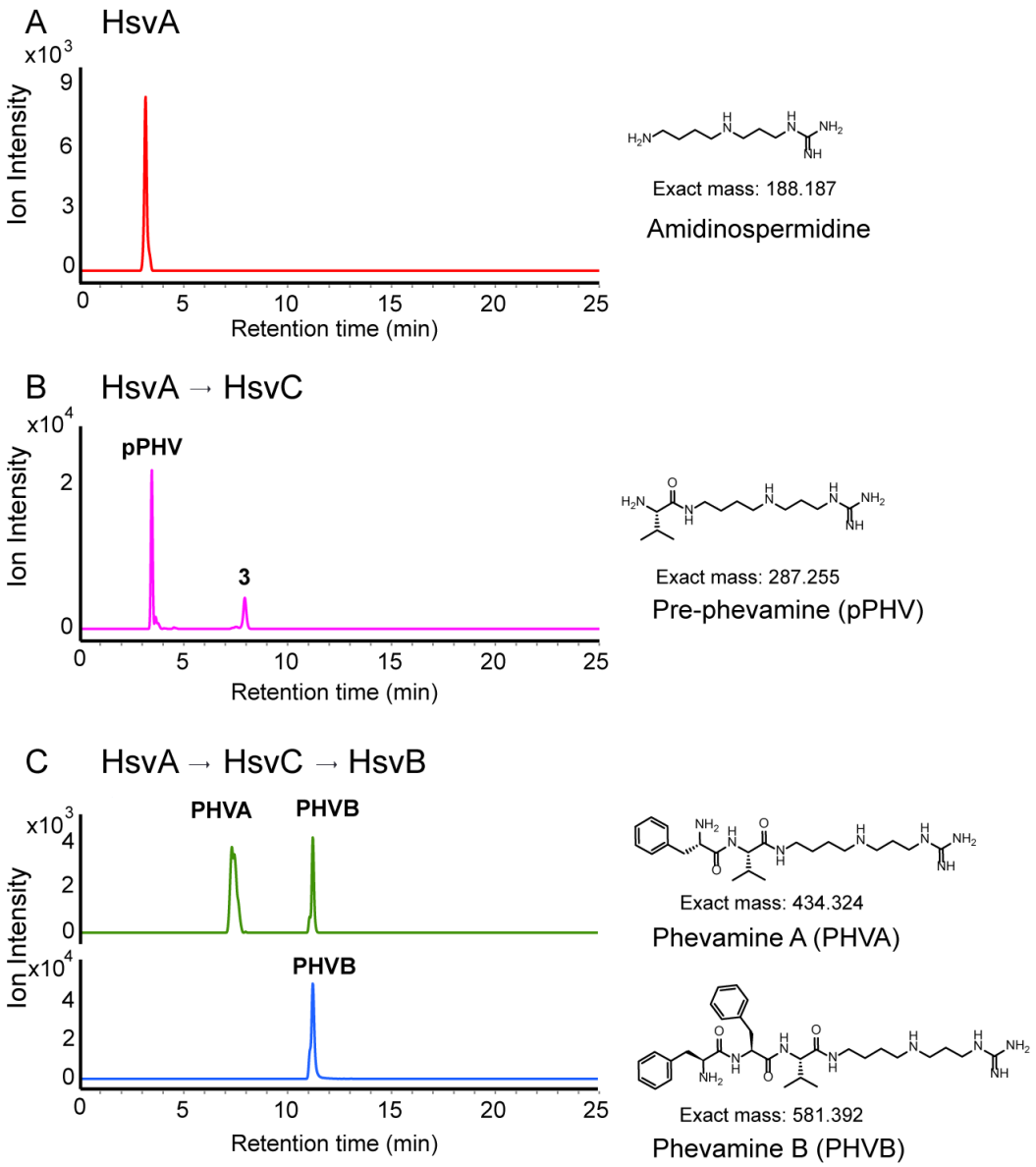


Figure S2.10. Production of phevamines and detection of pathway intermediates by sequential *in vitro* enzymatic assays. Extracted ion chromatograms are shown for *(A)* amidinospermidine (*m/z* 188.187) from HsvA reaction, *(B)* pre-phevamine (*m/z* 287.255) from HsvA reaction followed by HsvC reaction, and *(C)* phevamine A (*m/z* 434.324, green) and phevamine B (*m/z* 581.392, blue) from sequential reactions of HsvA, HsvC, and HsvB. In (*B*), the second small peak for pre-phevamine comes from in-source fragmentation of Compound 3, the mass of which corresponds to Val-Val-Val-amidinospermidine, an *in vitro* byproduct of the HsvC reaction. In (*C*), the second peak for phevamine A is the result of in-source fragmentation of phevamine B.

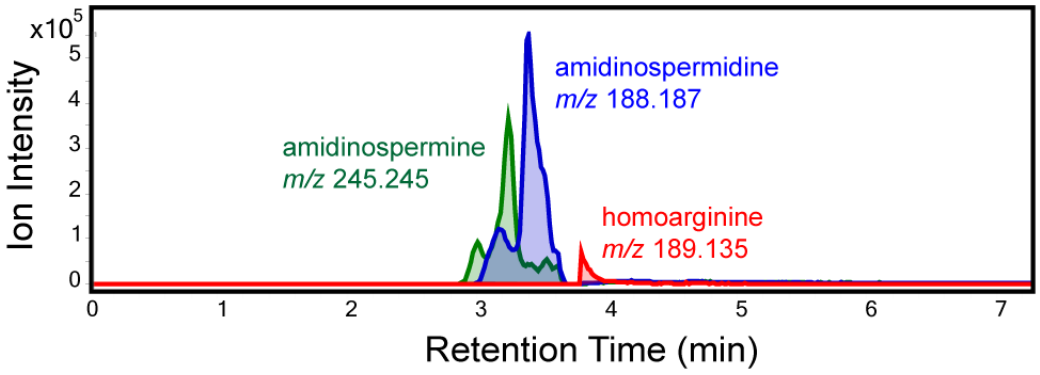


Figure S2.11. In vitro reconstitution of HsvA activity. HsvA was incubated with spermine, spermidine, and L-lysine as the acceptor of the amidino group and L-arginine as the donor of the amidino group. Extracted ion chromatograms are shown for amidinospermidine (*m/z* 188.187), amidinospermine (*m/z* 245.245), and homoarginine (*m/z* 189.135), which are the amidinotransfer products catalyzed by HsvA for spermidine, spermine, and arginine, respectively.

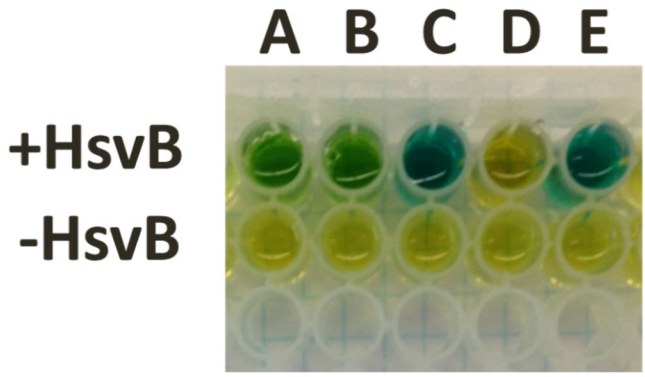


Figure S2.12. HsvB activates L-Phe as the preferred substrate. Several pairs of amino acids were incubated with HsvB and the released inorganic phosphate (P_i) was measured using the P_iColorLock assay described in the Methods. The amino acid pairs are (A) L-Ser and L-Orn, (B) L-Ala and L-Lys, (C) L-Ala and L-Phe, (D) L-Arg and L-Trp, and (E) 2x L-Phe. The assays containing L-Phe ((C) and (E)) produced the most P_i, indicating that L-Phe is the preferred substrate for HsvB.

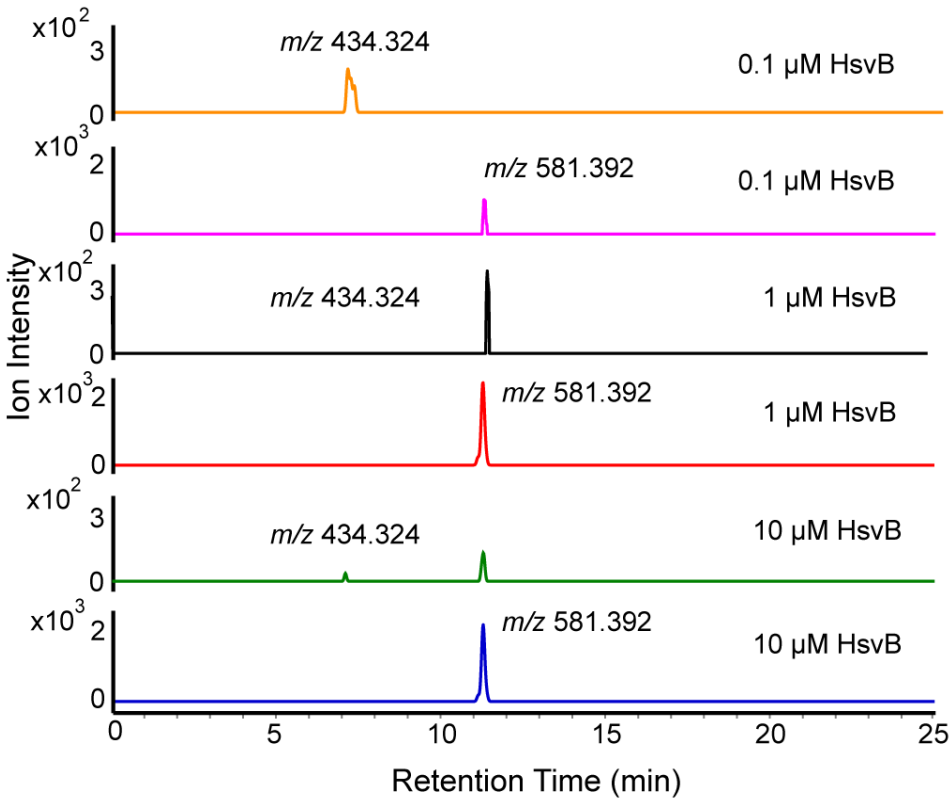
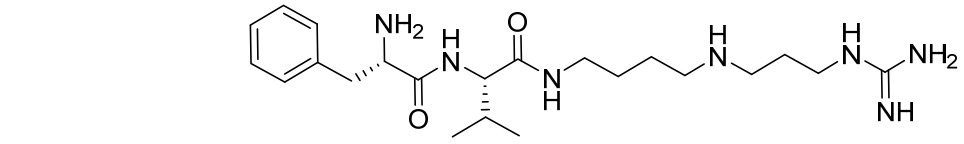


Figure S2.13. Production of phevamine A and B depends on HsvB concentration. Reaction mixtures containing 100 μM pPHV, 20 μM Phe, and varying amounts of HsvB were incubated at room temperature for 2 hours. As the concentration of HsvB increased, the amount of phevamine B (*m/z* 581.392) increased (pink, red, and blue). The ratio of phevamine A (*m/z* 434.324) (orange, black and green) to phevamine B is the highest at low concentration of HsvB. The peak around 11 minute retention time in phevamine A traces is a result of in-source fragmentation of phevamine B.



Phevamine A

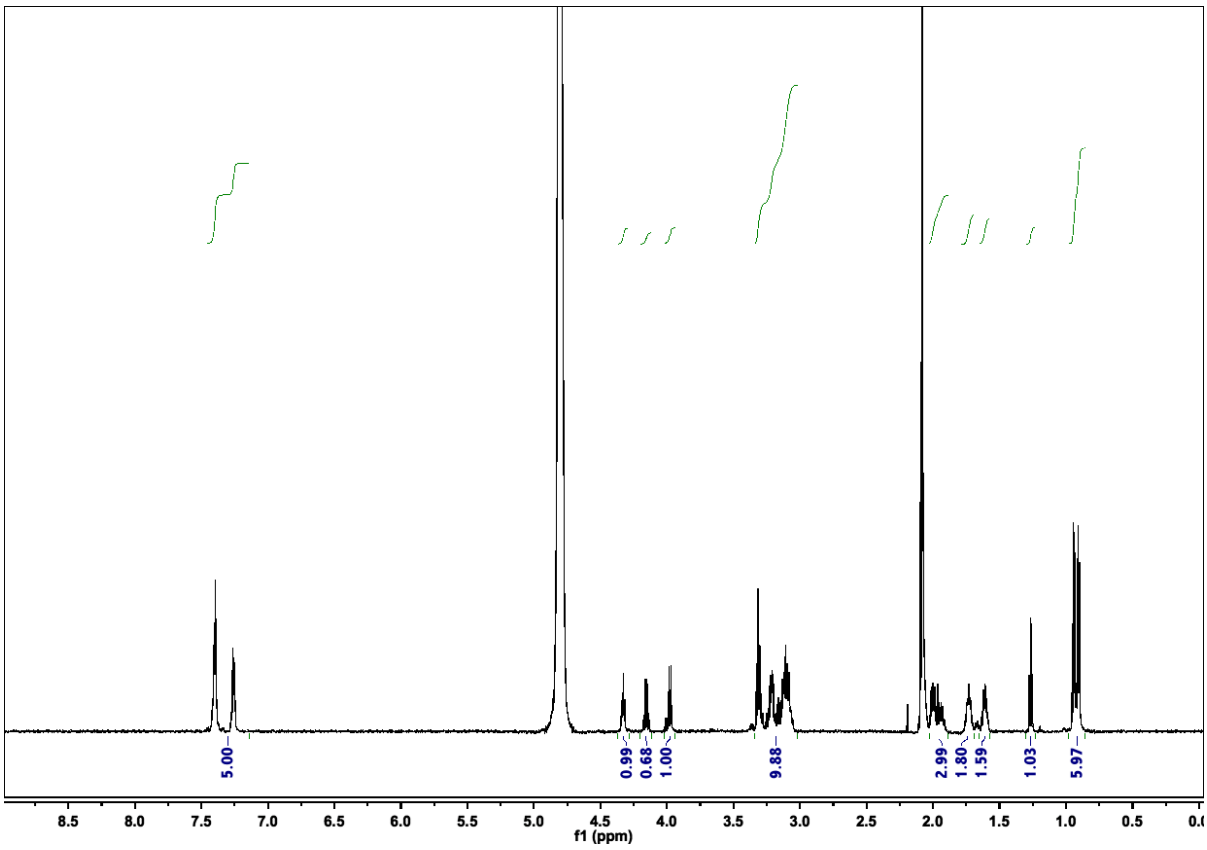


Figure S2.14. ^{1}H NMR (600 MHz) of *in vitro* synthesized and purified phevamine A. Spectrum acquired in D_2O.

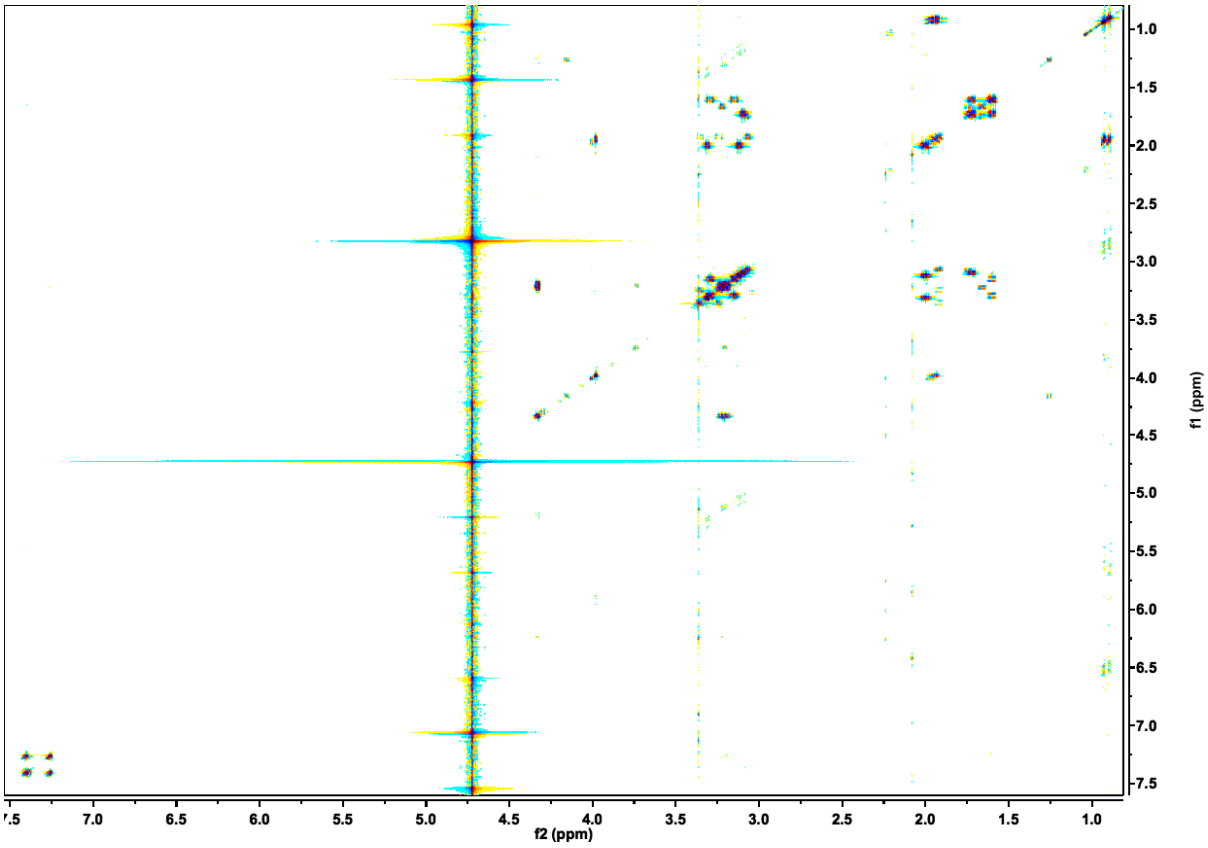


Figure S2.15. (^{1}H,^{1}H)-dqfCOSY spectrum (800 MHz) of *in vitro* synthesized and purified phevamine A. Spectrum acquired in CD_3OD.

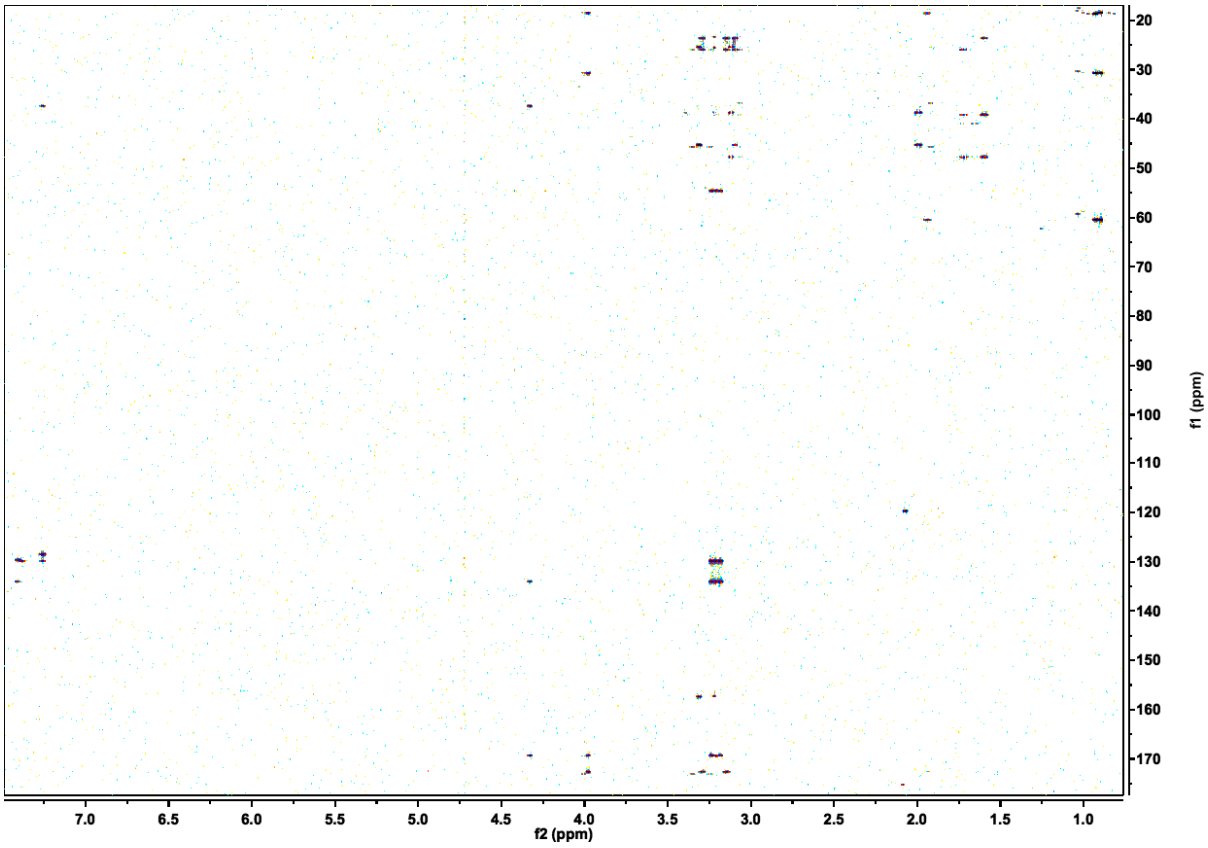


Figure S2.16. (^{1}H,^{13}C)-HMBC spectrum (800 MHz) of *in vitro* synthesized and purified phevamine A. Spectrum acquired in CD_3OD.

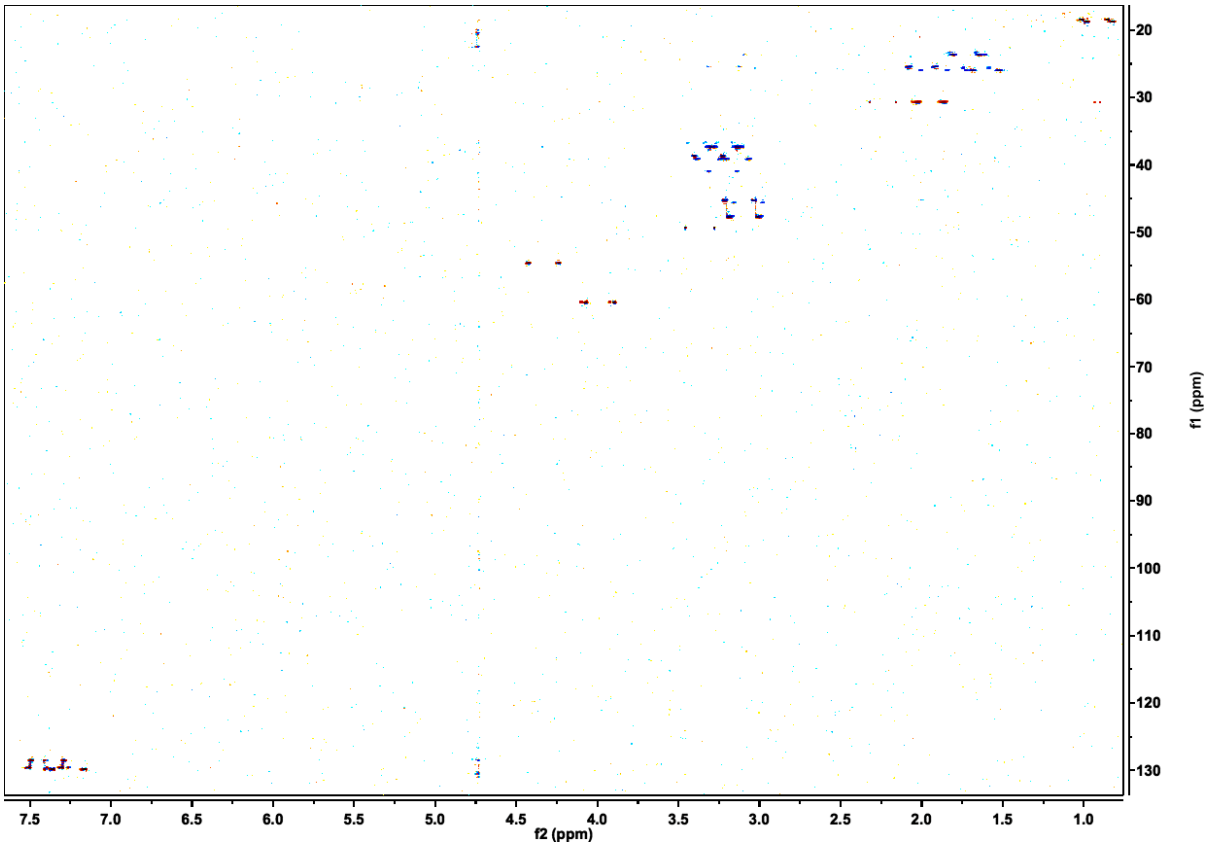
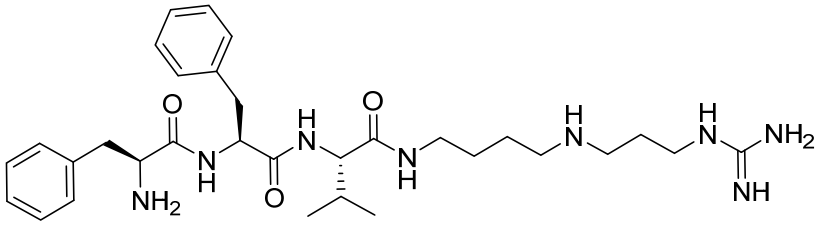


Figure S2.17. (^{1}H,^{13}C)-coupled-HSQC spectrum (800 MHz) of *in vitro* synthesized and purified phevamine A. Spectrum acquired in CD_3OD.



Phevamine B

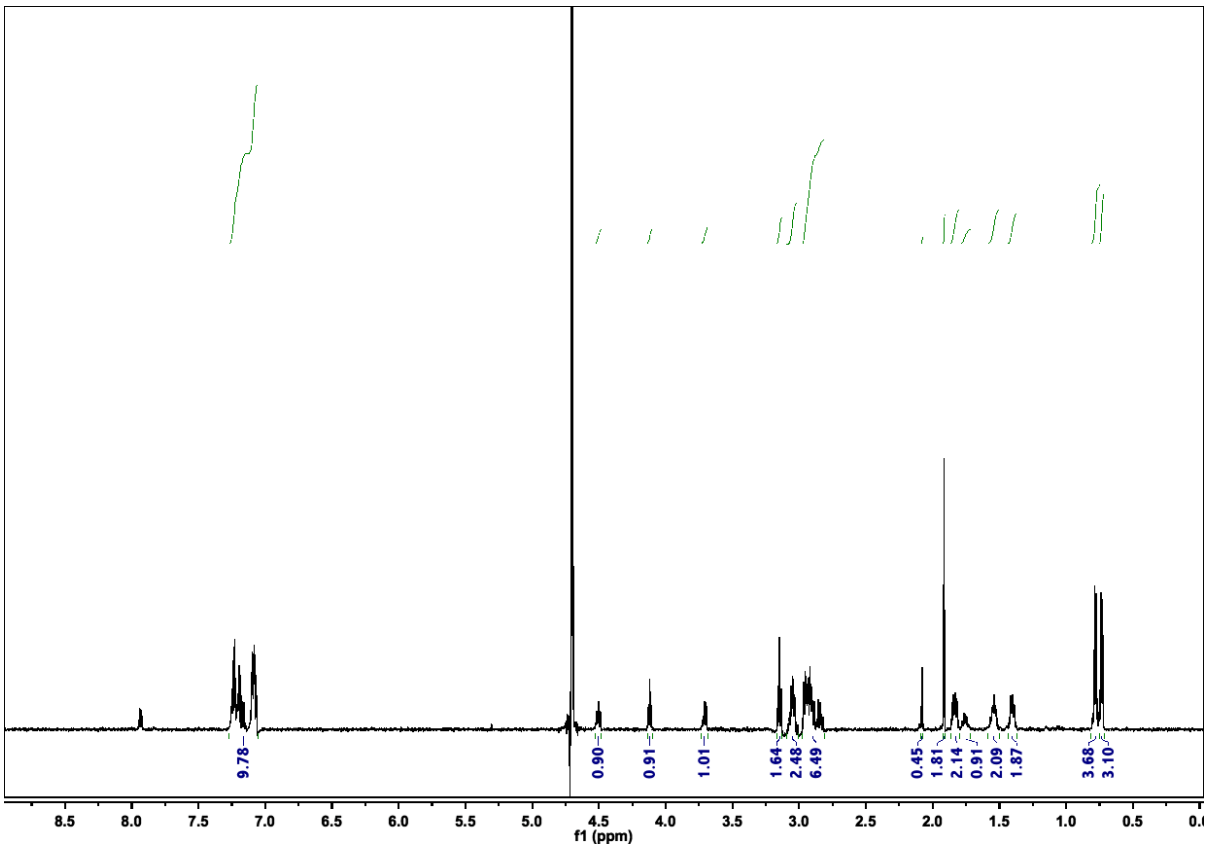


Figure S2.18. ^{1}H NMR (600 MHz) of *in vitro* synthesized and purified phevamine B. Spectrum acquired in CD_3OD.

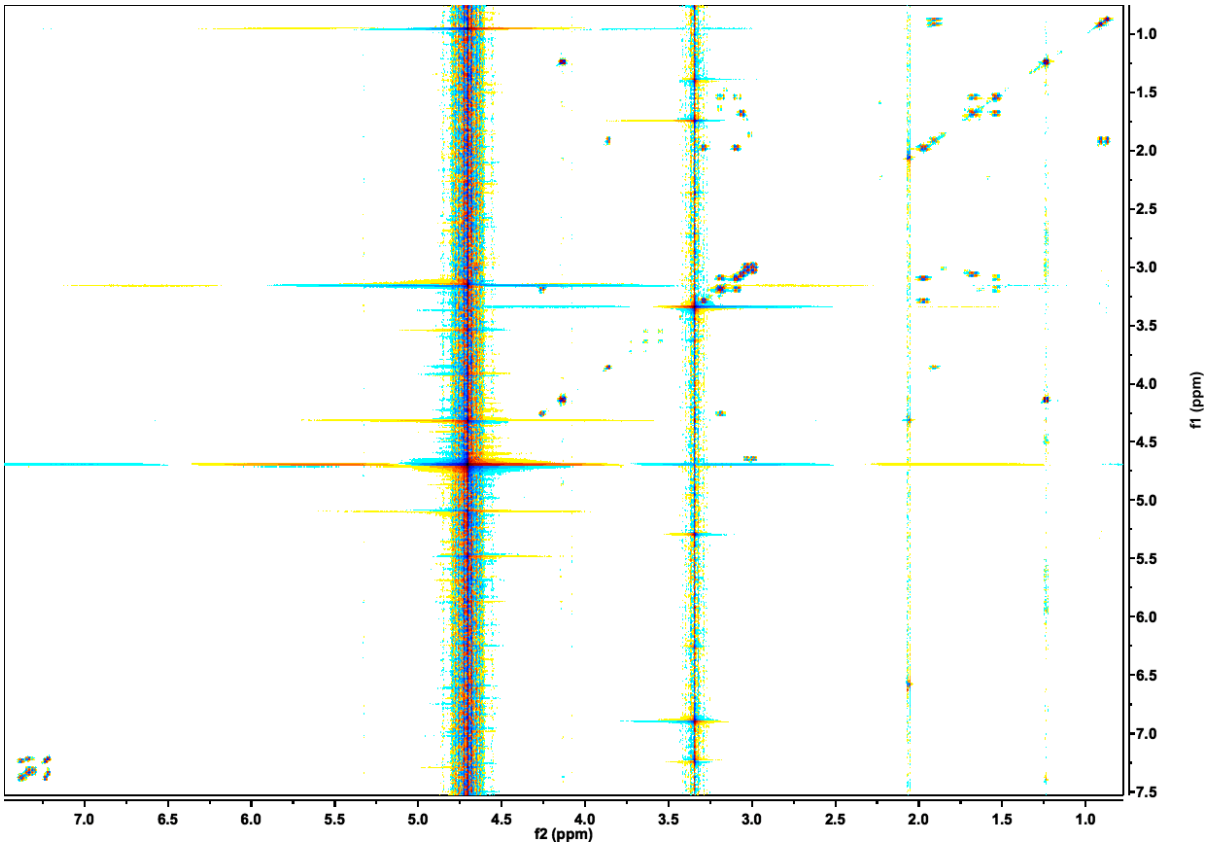


Figure S2.19. (^{1}H,^{1}H)-dqfCOSY spectrum (800 MHz) of *in vitro* synthesized and purified phevamine B. Spectrum acquired in CD_3OD.

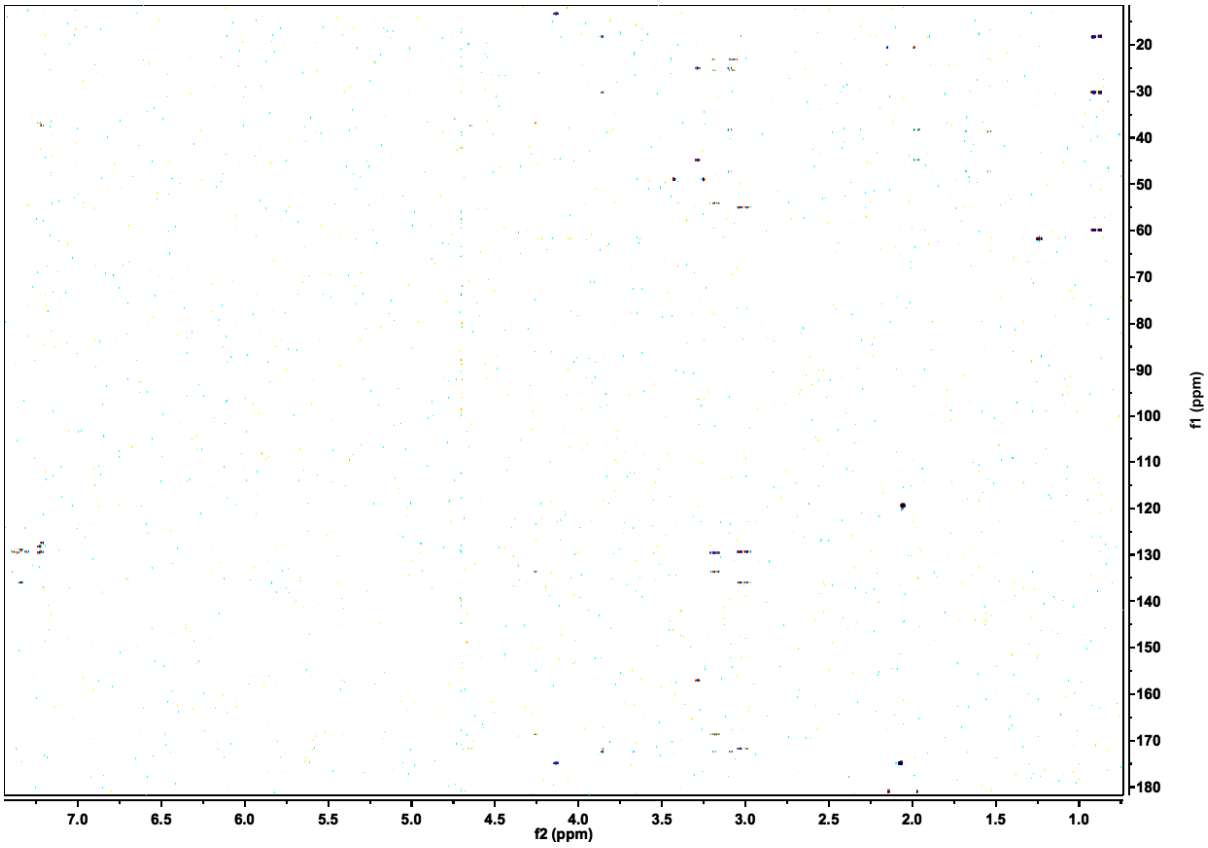


Figure S2.20. (^{1}H,^{13}C)-HMBC spectrum (800 MHz) of *in vitro* synthesized and purified phevamine B. Spectrum acquired in CD_3OD.

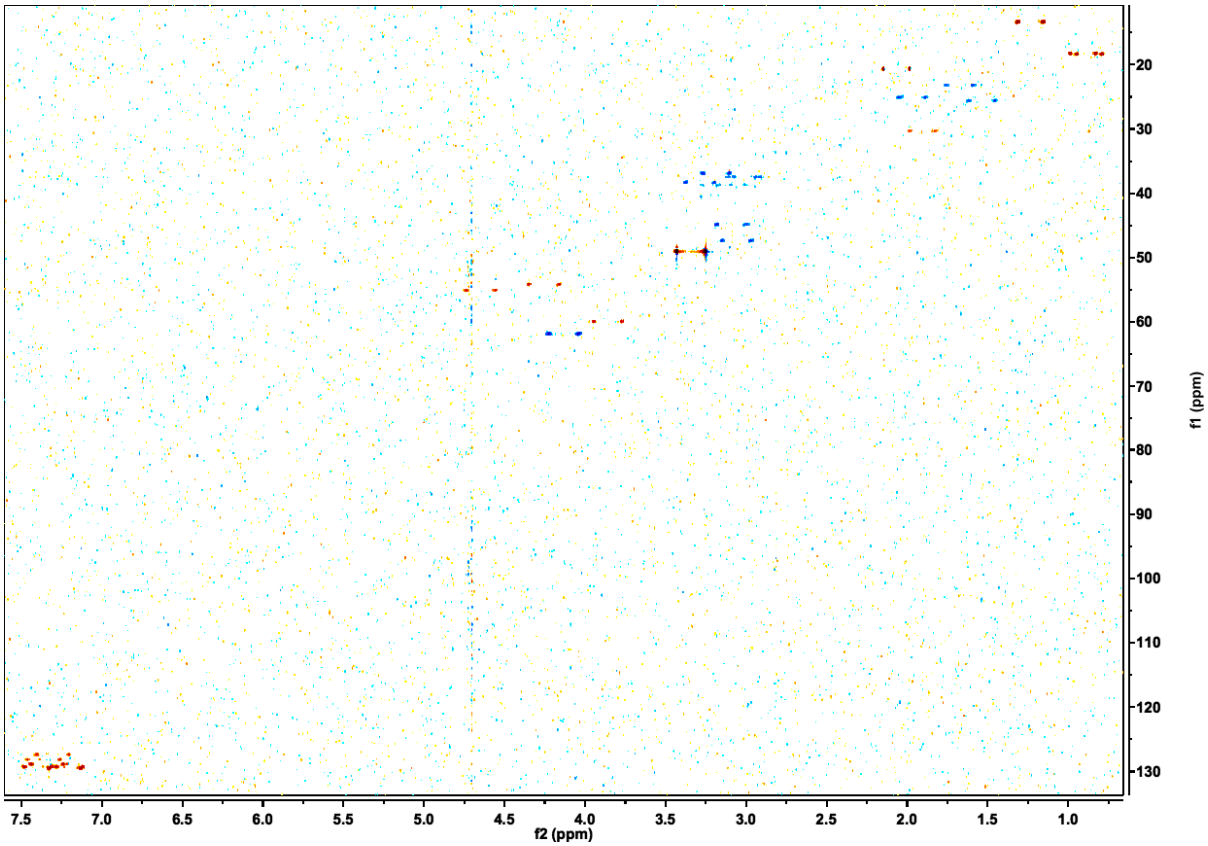


Figure S2.21. (^{1}H,^{13}C)-coupled-HSQC spectrum (800 MHz) of *in vitro* synthesized and purified phevamine B. Spectrum acquired in CD_3OD.

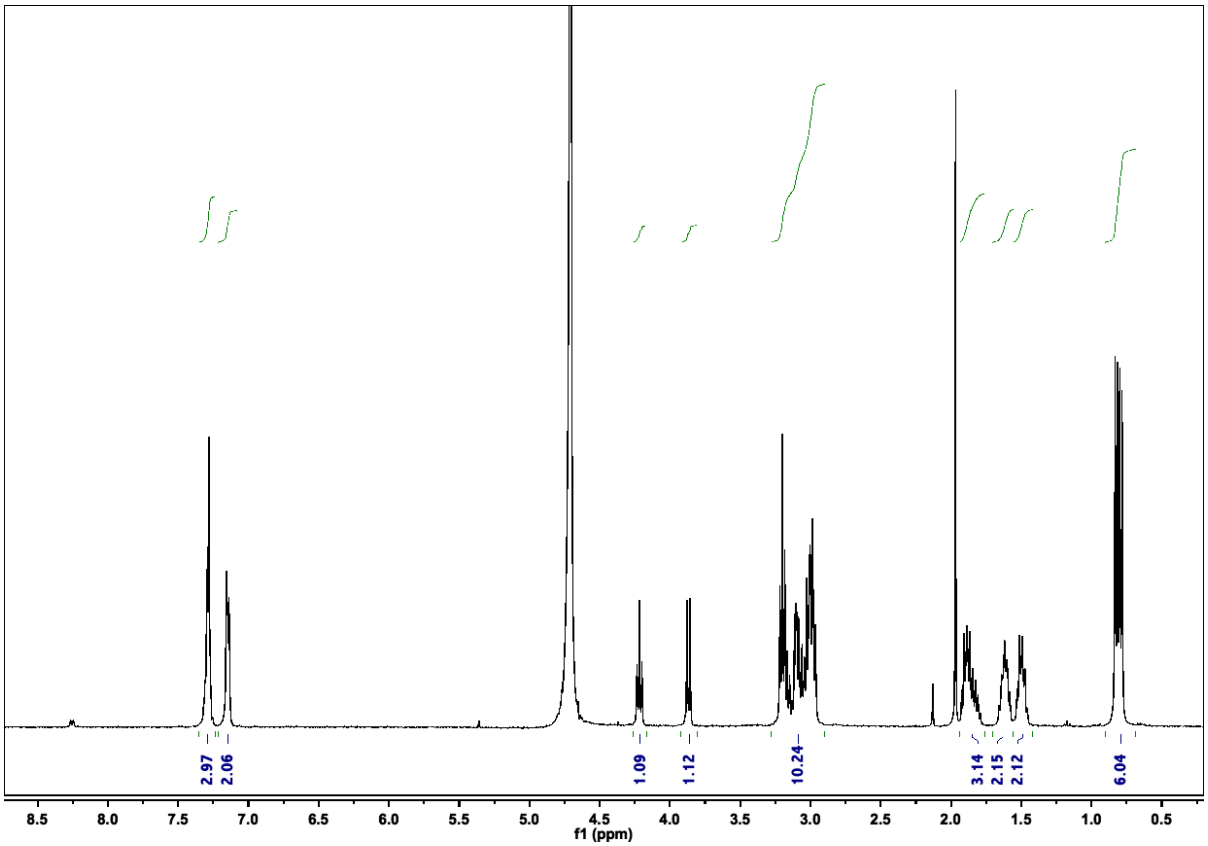


Figure S2.22. ^{1}H NMR (400 MHz) of the phevamine A synthetic standard (9). Spectrum acquired in D_2O.

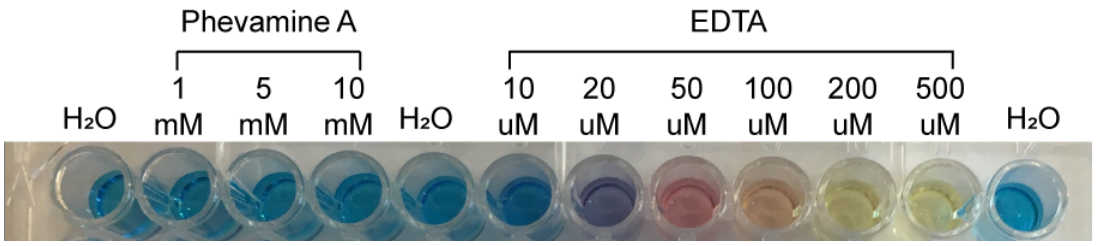


Figure S2.23. Iron binding activity of phevamine A. The ability of phevamine A to bind iron was tested and compared to EDTA standards using a chromeazurol S (CAS) assay. Upon binding of ferric iron, the chromeazurol S dye changes color from blue to yellow. No color change was observed for phevamine A at a concentration of up to 10 mM, thus it is unlikely that phevamine A acts as a siderophore to bind ferric iron.

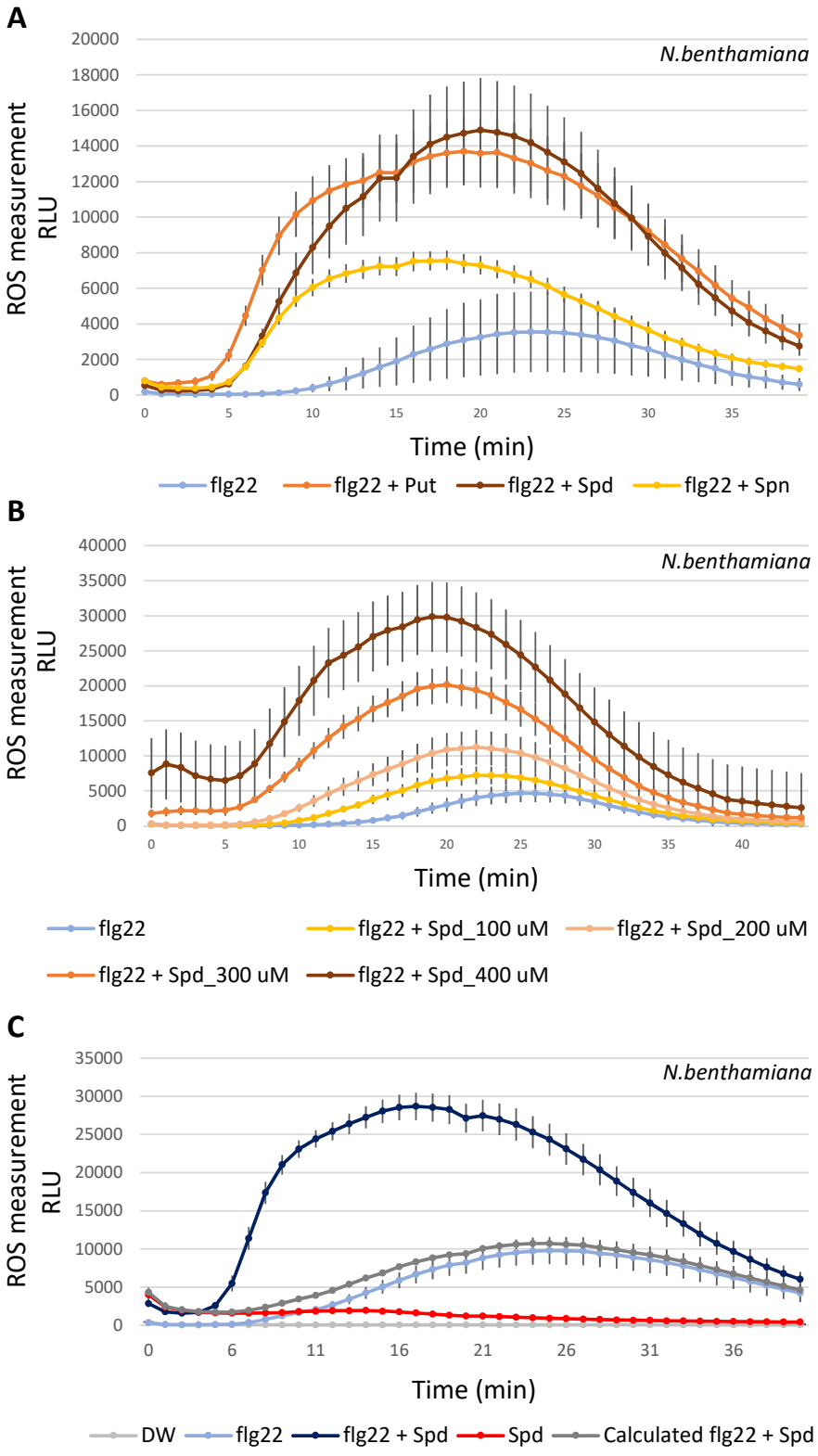


Figure S2.24. Polyamines enhance the flg22-induced ROS burst in *N. benthamiana*. *(A)* Various polyamines potentiate the flg22-induced ROS burst. Leaf disks were challenged with

10 nM of flg22 with or without polyamines at 400 μM. Put: putrescine, Spd: spermidine, and Spn: spermine. *(B)* Dose-dependent effect of spermidine on the flg22-induced ROS burst. Leaf disks were challenged with 10 of nM flg22 and various concentrations of spermidine. *(C)* The increase of the flg22-induced ROS burst by spermidine is not due to an additive effect. Leaf disks were treated with distilled water (DW), 10 nM of flg22, 10 nM of flg22 with 400 μM of spermidine, or 400 μM of spermidine only. Experiments presented in panels (*A*) and (*C*) were performed more than 3 times with similar results. The experiment presented in panel (*B*) was done twice with similar results. Error bars represent standard errors.

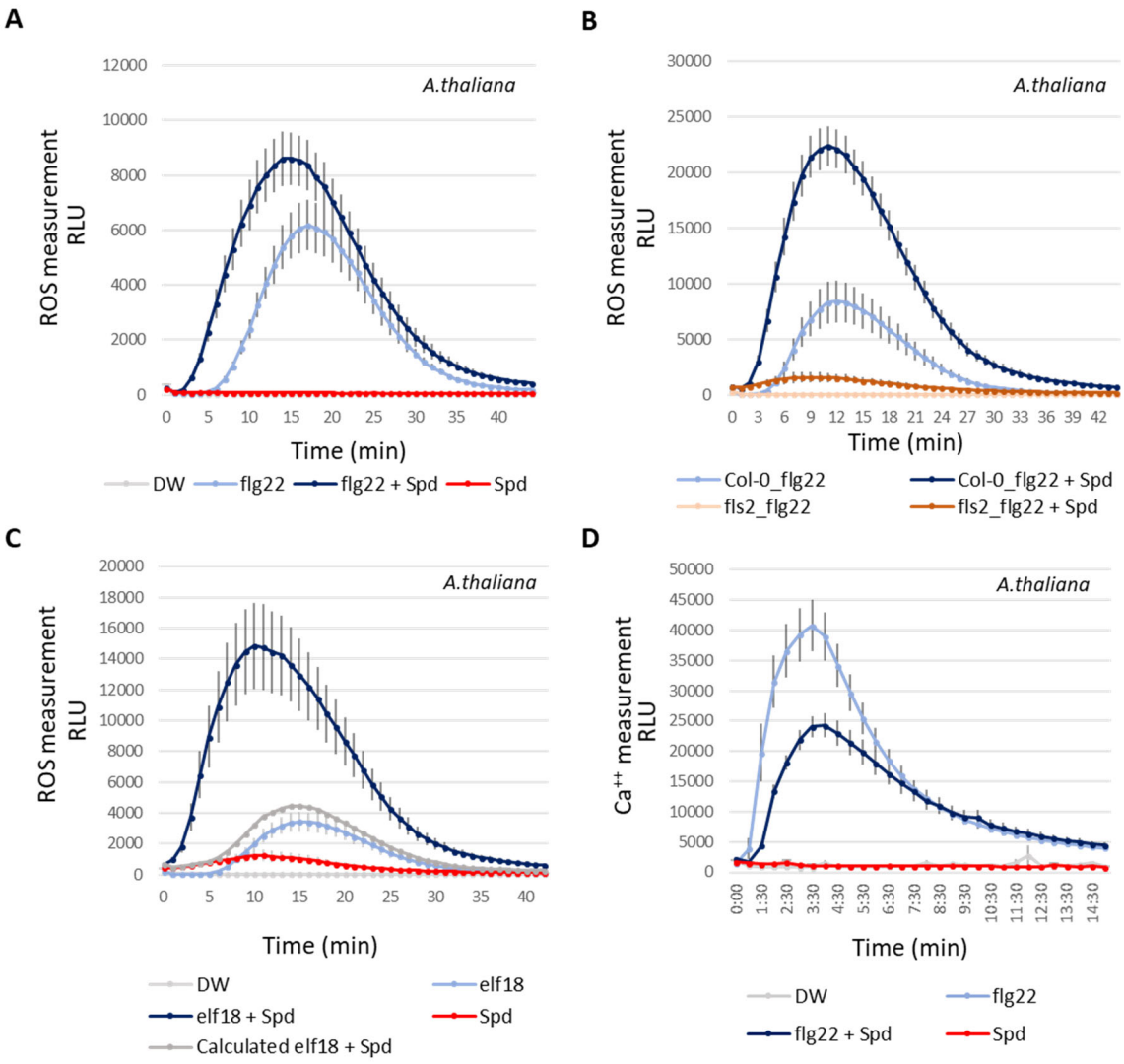


Figure S2.25. Effect of spermidine on early MTI responses in *A. thaliana*. (*A*) Spermidine potentiates the flg22-induced ROS burst. Leaf disks were treated with distilled water (DW), flg22 at 10 nM, flg22 with spermidine at 300 µM or spermidine only at 300 µM. ROS levels were measured. This assay was repeated more than 3 times with similar results. (*B*) Potentiation of the flg22 ROS burst by spermidine is dependent on FLS2. Leaf disks from Col-0 or the *fls2* mutant were challenged with flg22 only at 20 nM or with spermidine at 300 µM. This experiment was performed twice with similar results. (*C*) Spermidine potentiates the elf18-induced ROS burst. Leaf disks were challenged with DW, elf18 at 50 nM, elf18 supplemented with spermidine at 400 µM, or with spermidine only at 400 µM. This experiment was performed twice. (*D*) Spermidine suppresses the flg22-induced calcium burst. Col-

0_pMAQ2 leaf disks were treated with DW, flg22 at 25 nM, flg22 at 25 nM with spermidine at 400 μM, or with spermidine only. This experiment was repeated more than 3 times. Error bars represent standard errors.

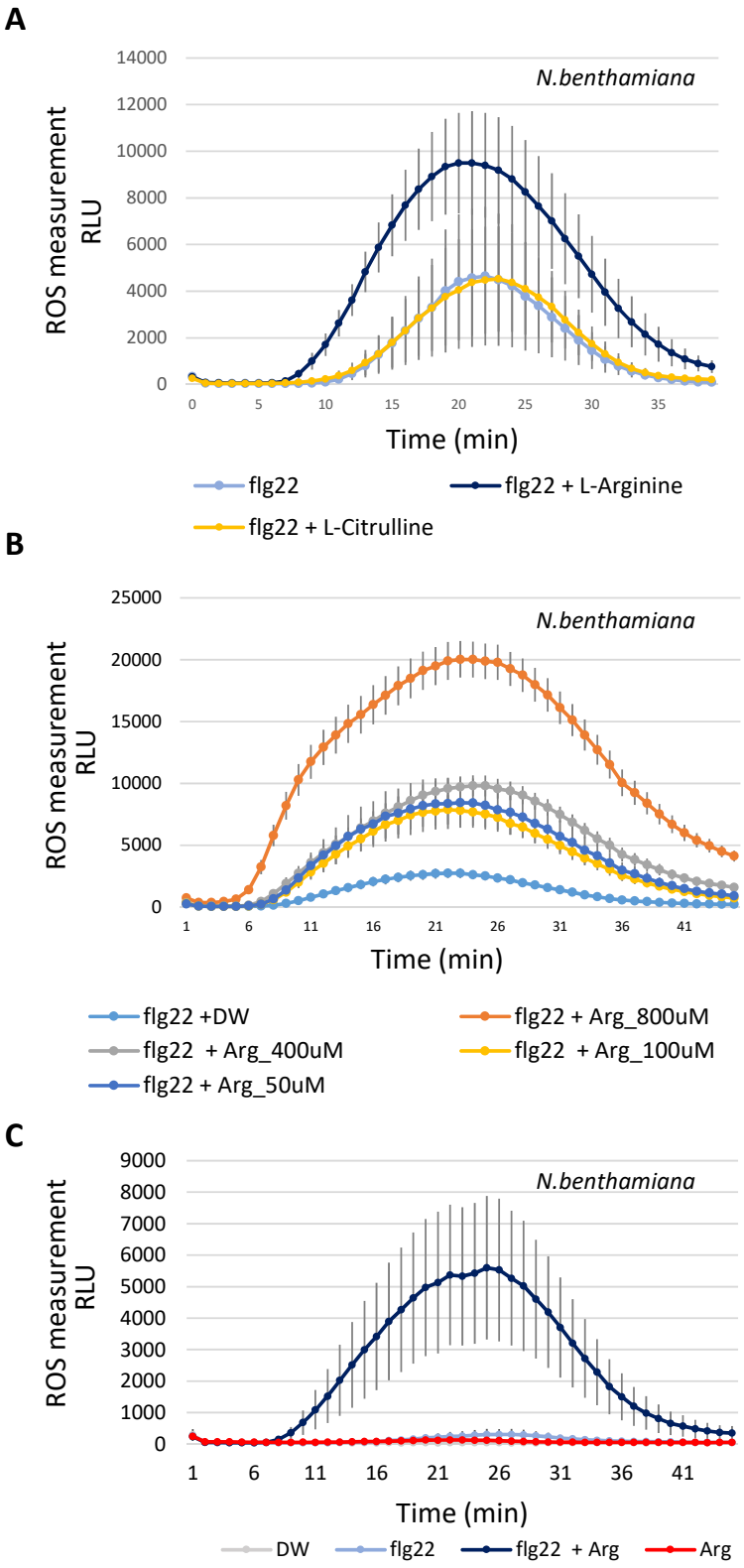


Figure S2.26. Arginine potentiates the flg22-induced ROS burst in *N. benthamiana*. *(A)* L-Arginine potentiates the flg22-induced ROS burst, but L-citrulline does not. flg22 was used

at 10 nM, arginine and citrulline were used at 400 µM. This experiment was performed twice with similar results. *(B)* Dose-dependent effect of arginine on the flg22-induced ROS burst. flg22 was used at 100 nM and arginine at various concentrations. This experiment was performed twice with similar results. *(C)* The increase of the flg22-induced ROS burst by arginine is not due to an additive effect. flg22 was used at 10 nM and arginine at 400 µM. This experiment was repeated at least 3 times with similar results. Error bars represent standard errors.

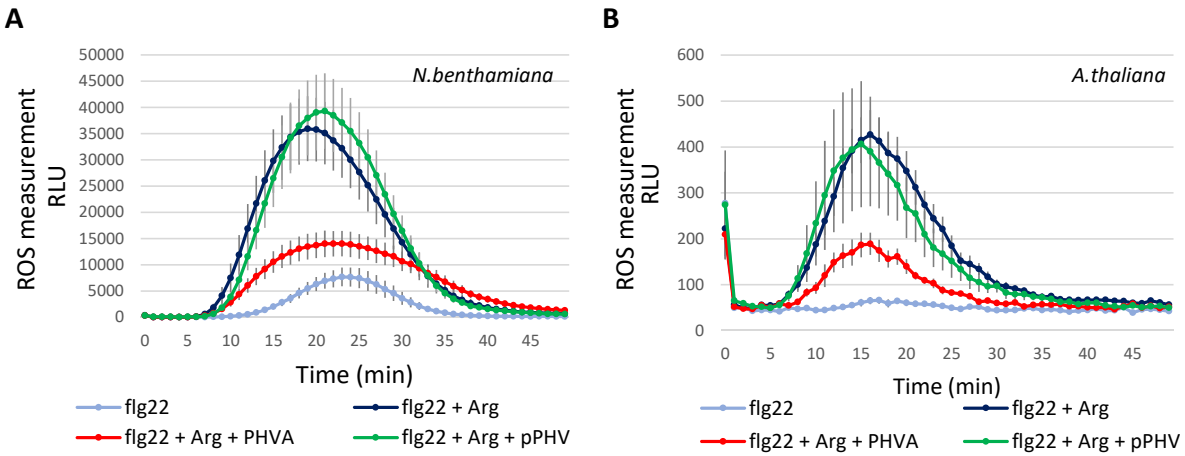


Figure S2.27. Phevamine A, but not pre-phevamine, suppresses the arginine potentiation of the flg22-induced ROS burst in both *N. benthamiana* and *A. thaliana*. *(A)* Phevamine A, but not pre-phevamine, suppresses the arginine ROS potentiation in *N. benthamiana.* flg22 was used at 5 nM; arginine, pre-phevamine and phevamine A at 300 µM. *(B)* Phevamine A, but not pre-phevamine, suppresses the arginine ROS potentiation in *A. thaliana.* flg22 was used at 5 nM; arginine, pre-phevamine, and phevamine A at 300 µM. PHVA: Phevamine A, pPHV: pre-phevamine. Experiments are representatives of 3 independent repeats with similar results.

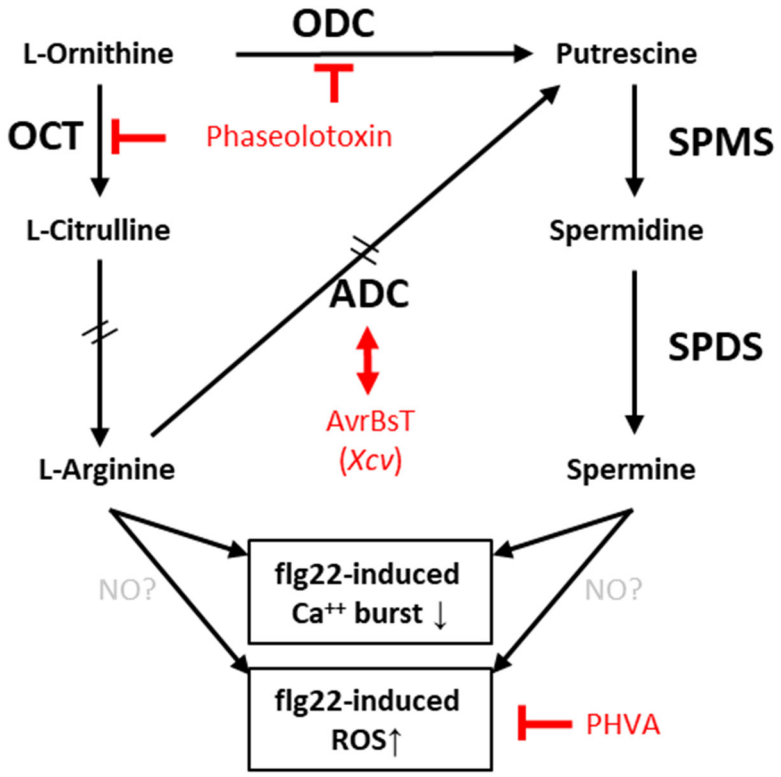


Figure S2.28. Arginine and polyamine biosynthetic pathways are targeted by pathogens via different mechanisms. Spermidine and L-arginine potentiate the flg22-induced ROS burst. The role of nitric oxide production in this potentiation remains uncertain.[128] Phevamine A suppresses the spermidine/L-arginine potentiation of the flg22-induced ROS burst. Phaseolotoxin inhibits the synthesis of arginine and polyamines pathway by suppressing the activity of ODC and OCT, while phevamine A targets the potentiation mediated by spermidine and L-arginine. The effector AvrBsT from *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) interacts (red double arrow) with ADC1 in pepper.[129] ODC: ornithine decarboxylase, OCT: ornithine carbamoyltransferase, ADC: arginine decarboxylase, SPDS: spermidine synthase, SPMS: spermine synthase, ROS: reactive oxygen species, NO: nitric oxide, PHVA: phevamine A. Double dashed arrows represent pathways with intermediates not shown. Adapted from [77].

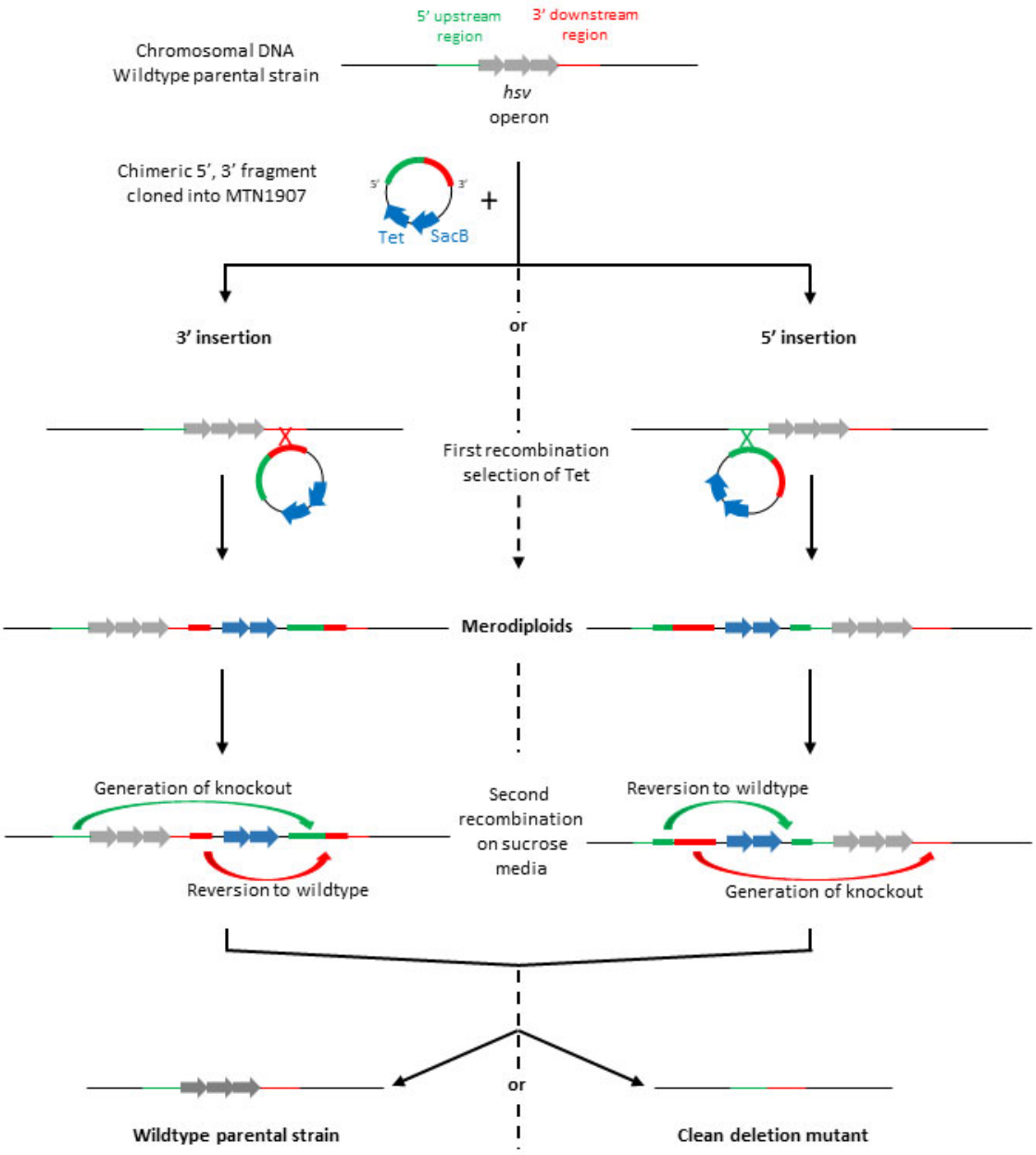


Figure S2.29. Strategy to generate *P. syringae* clean deletion mutants. A fusion of the 5' and 3' flanking region of the *hsv* operon was generated by chimeric PCR and cloned into MTN1907. After transformation of the parental strains (*Pto*, *Pto*-Cor $^{-}$, *Pto*D28E), the first recombination integrates the plasmid either at the 5 'or 3' end of the *hsv* operon. Merodiploids are selected on media containing tetracycline and genotyped by PCR. The second recombination event either leads to a reversion to the original locus or the generation of the

clean mutant. Recombined strains were selected on plates containing sucrose (5%) and clean deletion mutants were identified by PCR (see Figure S2.30).

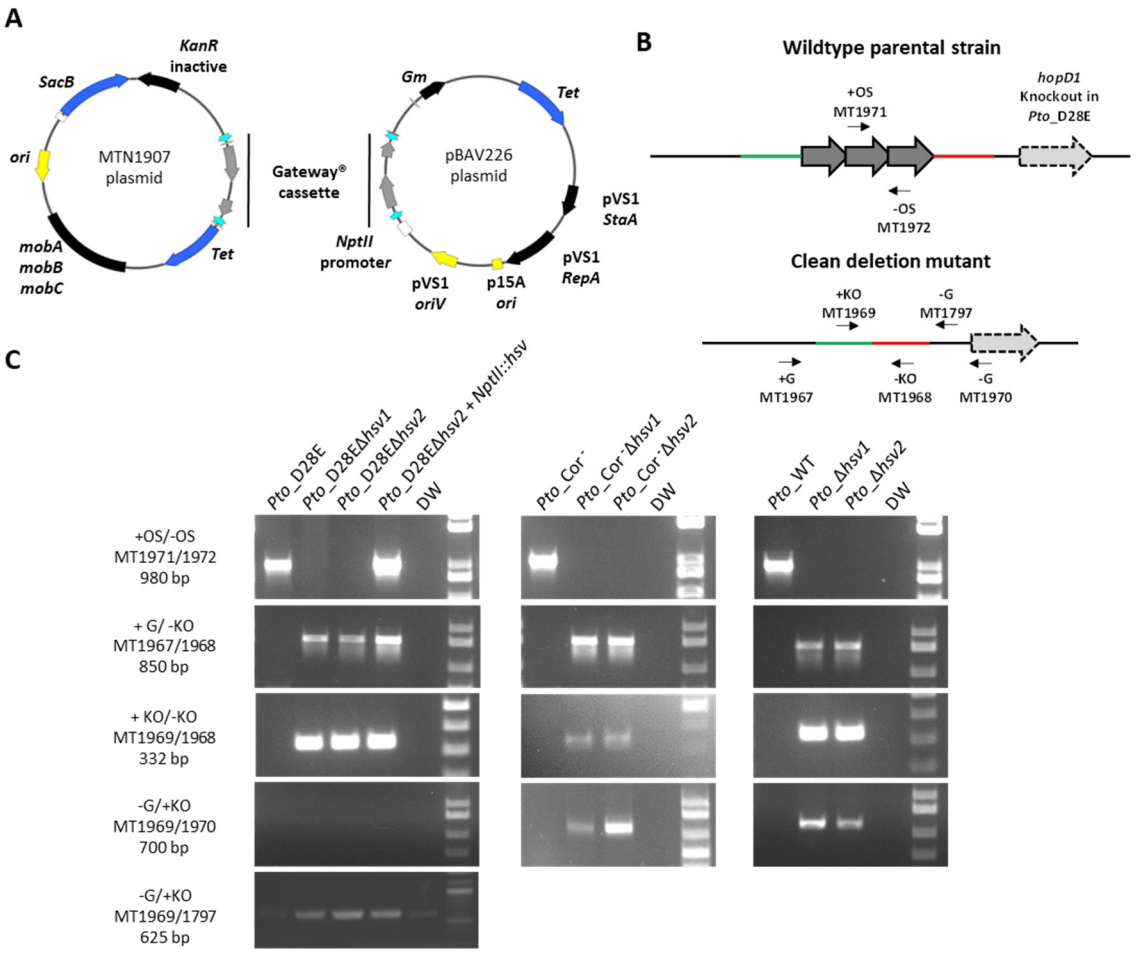
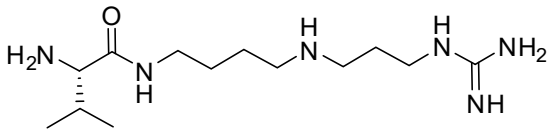


Figure S2.30. Genotyping of the clean deletion mutants and the complemented strain. *(A)* Graphical representation of the plasmids used for molecular manipulation of *P. syringae*. The left panel represents the MTN1907 plasmid used to engineer the mutants. The right panel represents the pBAV226 plasmid in which the *hsv* operon was cloned downstream of the *NptII* promoter for complementation of *Pto*D28E_Δ*hsv2*. Tet: confers resistance to tetracycline*; mobA, mobB, mobC*: RSF1010 mobilization genes; *ori*: high-copy-number origin of replication; *SacB*: a secreted levansucrase that renders bacteria sensitive to sucrose for counter-selection; *KanR*_inactive: aminoglycoside phosphotransferase gene inactive in *P. syringae*; pVS1 *StaA*: the stability protein from the *Pseudomonas* plasmid pVS1; pVS1 *RepA*: a replication protein; p15A *ori*: p15A origin of replication for propagation in *E. coli* cells; pVS1 *oriV*: origin of replication of the pVS1 plasmid; *Gm*: the gentamycin acetyltransferase. *(B)* Graphical representation of the *hsv* locus in the wildtype parental strains (*Pto*-WT, *Pto*-Cor $^{-}$,

*Pto*D28E), and the clean deletion mutants. Arrows represent the primers used for the genotyping. OS: operon specific primer, G: genomic primer, KO: knockout primer, +: forward primer, -: reverse primer (see Supplementary Table S2.2 for primer sequences). (*C*) Analytical PCRs performed to confirm the genotype of the mutants generated in this study. PCR programs include a 1 min extension time, designed for amplification in mutant strains, but too short for amplification in the wildtype parental strains with the exception of the PCRs using operon specific primers.



Pre-phevamine

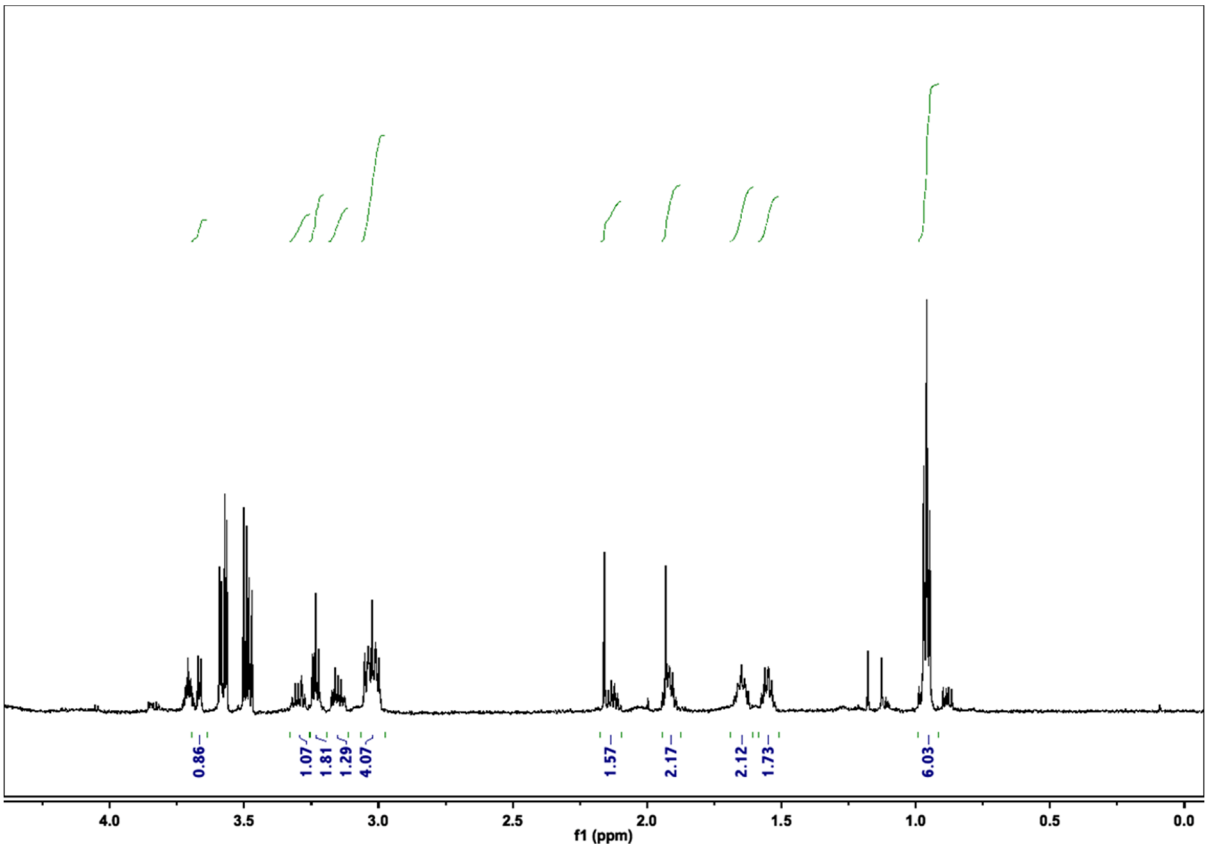


Figure S2.31. ^{1}H NMR (600 MHz) of *in vitro* purified pre-phevamine. Spectrum acquired in D_2O.

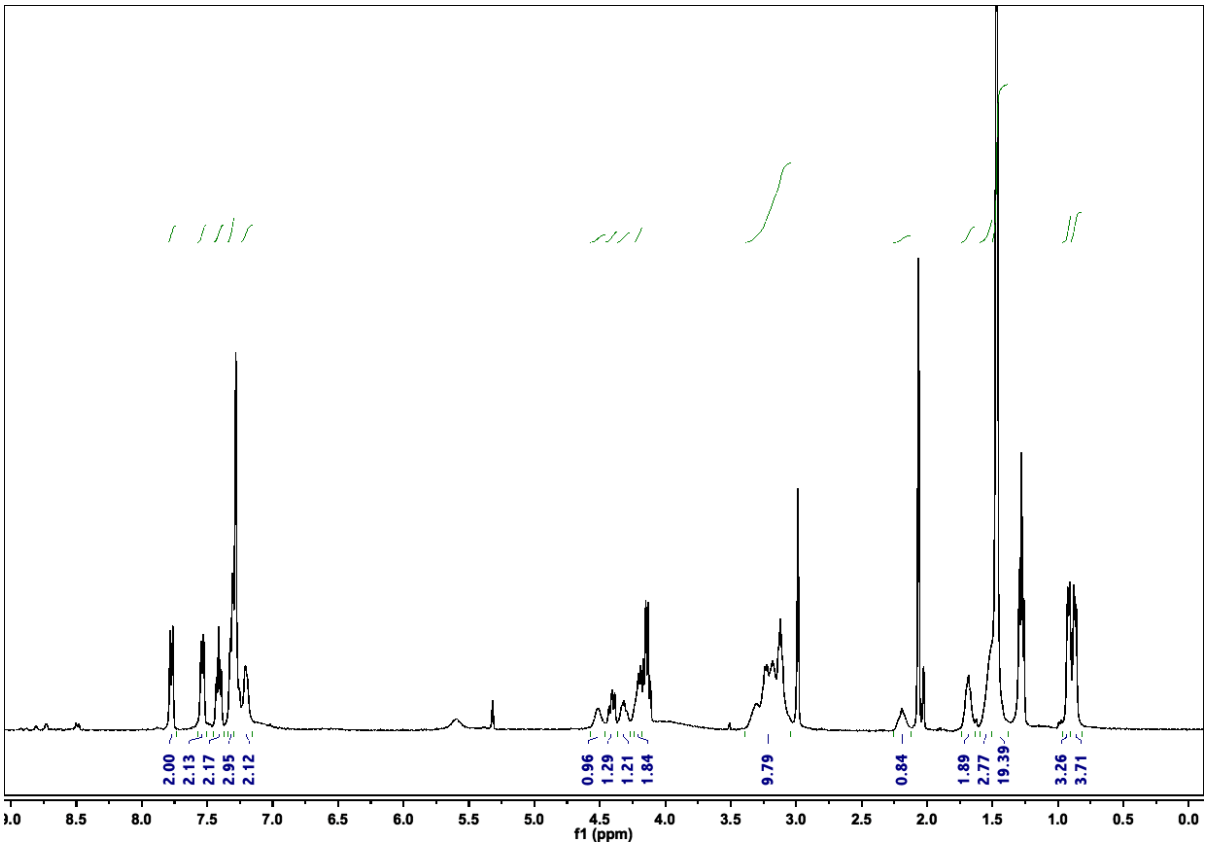


Figure S2.32. ^{1}H NMR (400 MHz) of Fmoc-(*S*)-Val-bis-Boc-spermidine (5). Spectrum acquired in $CDCl_3$.

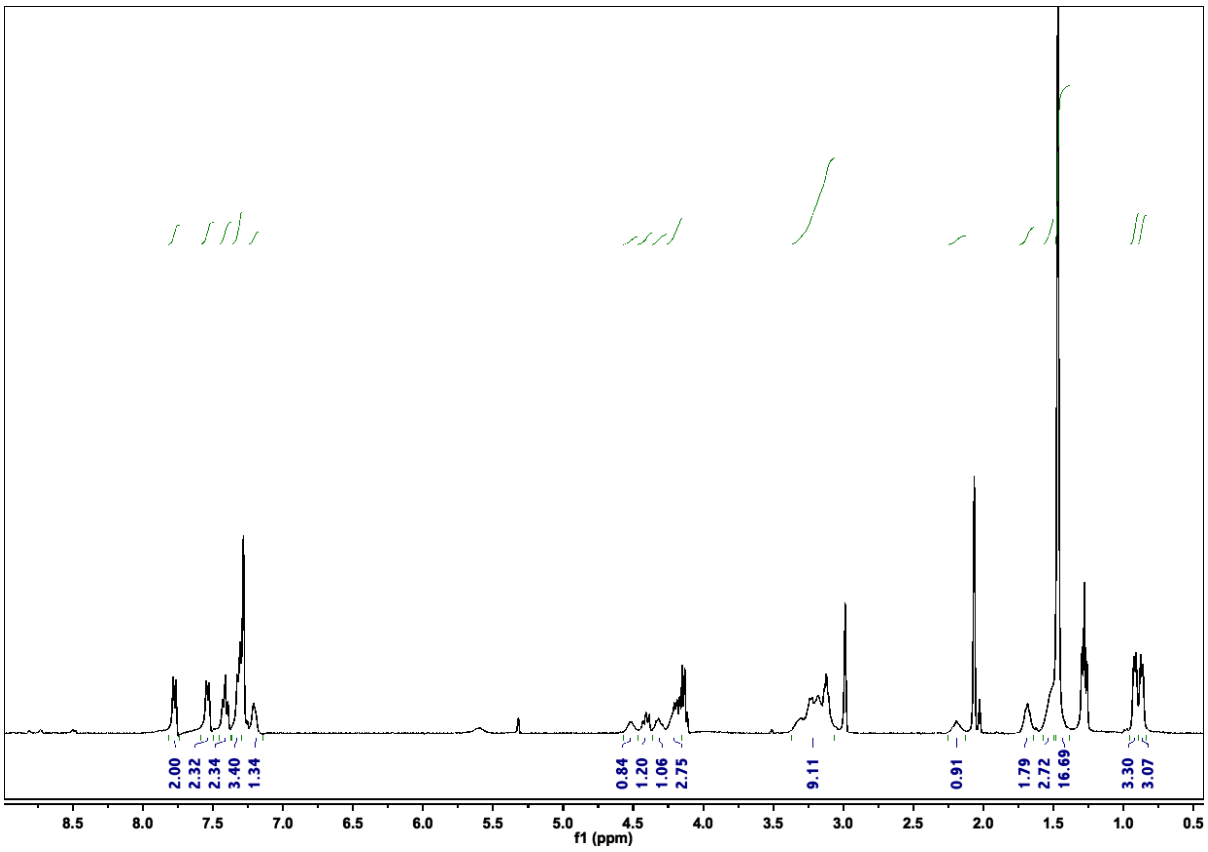


Figure S2.33. ^{1}H NMR (400 MHz) of Fmoc-(*S*)-Phe-(*S*)-Val-bis-Boc-spermidine (6). Spectrum acquired in $CDCl_3$.

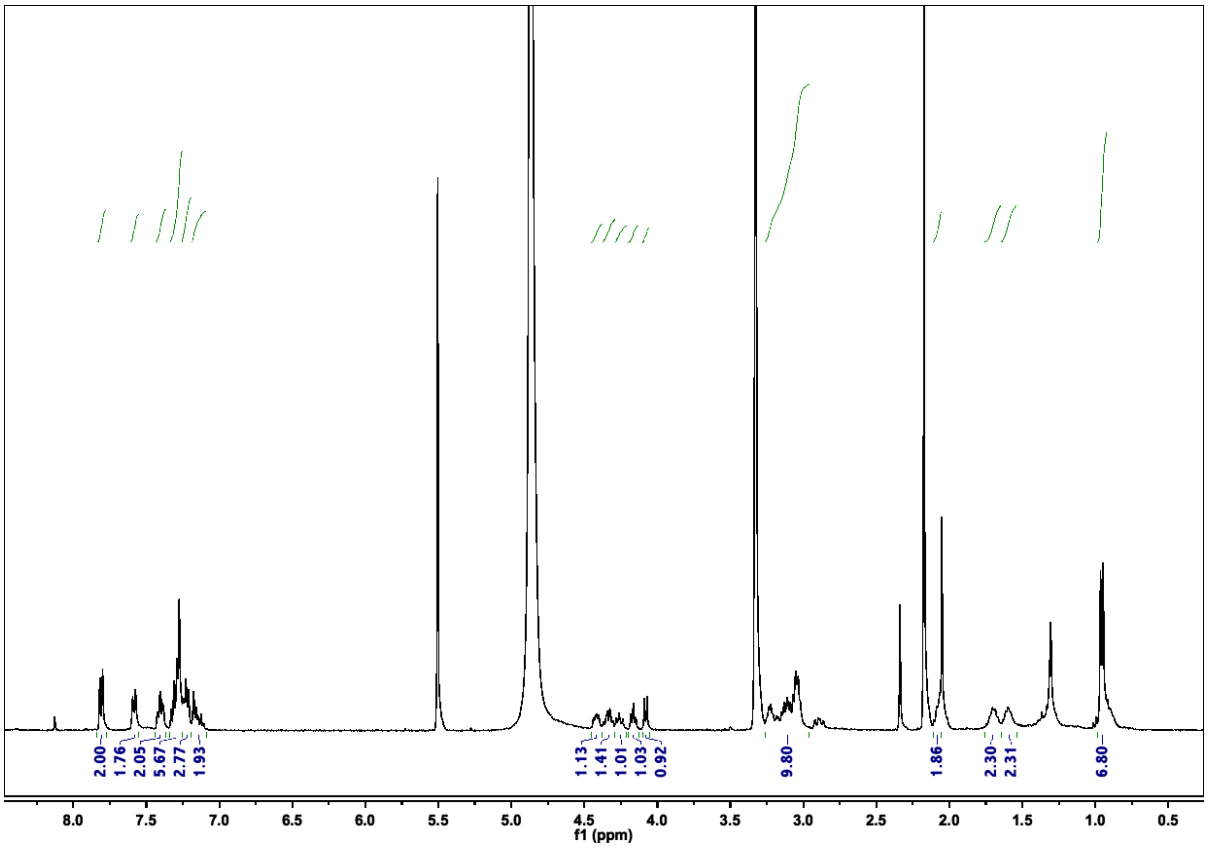


Figure S2.34. ^{1}H NMR (400 MHz) of Fmoc-(*S*)-Phe-(*S*)-Val-spermidine (7). Spectrum acquired in CD_3OD.

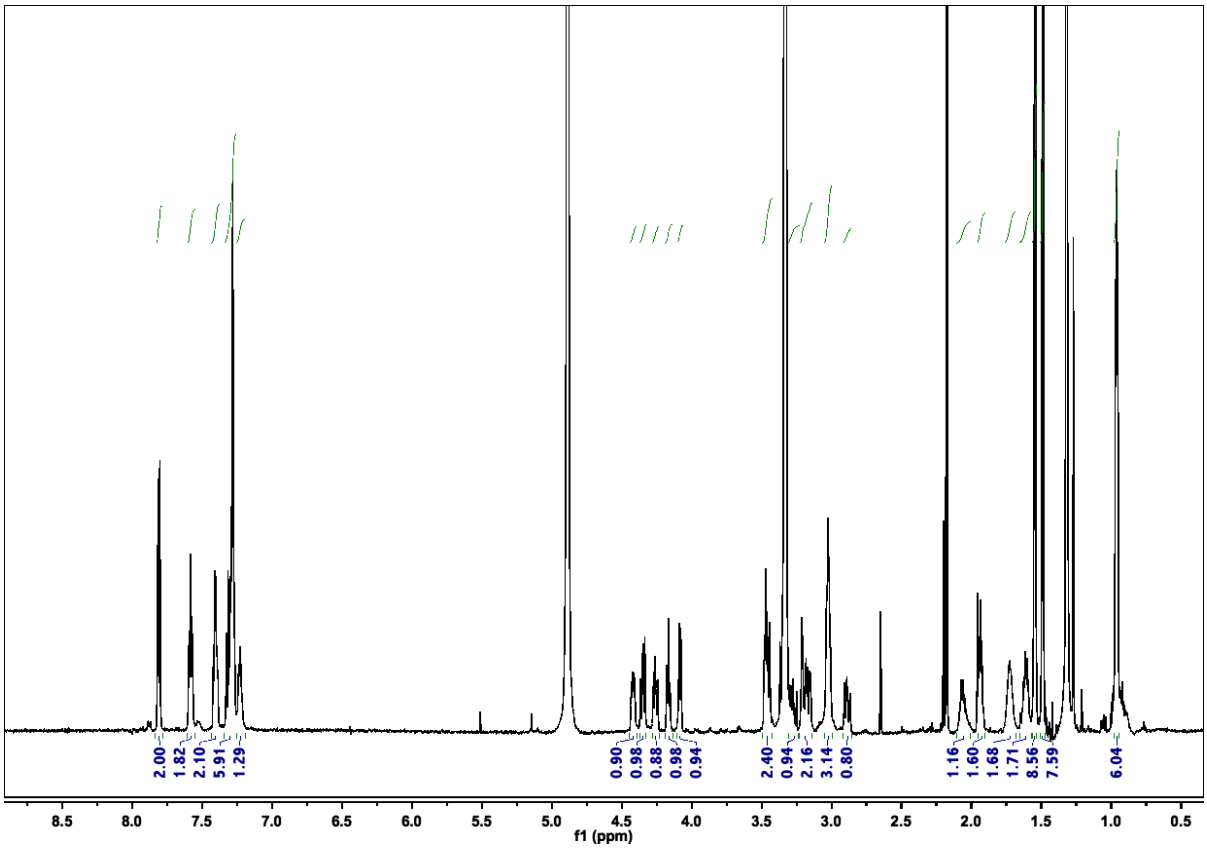


Figure S2.35. ^{1}H NMR (600 MHz) of Fmoc-(*S*)-Phe-(*S*)-Val-spermidine-*N*,*N*'-di-Boc-1*H*-guanidine (8). Spectrum acquired in CD_3OD.

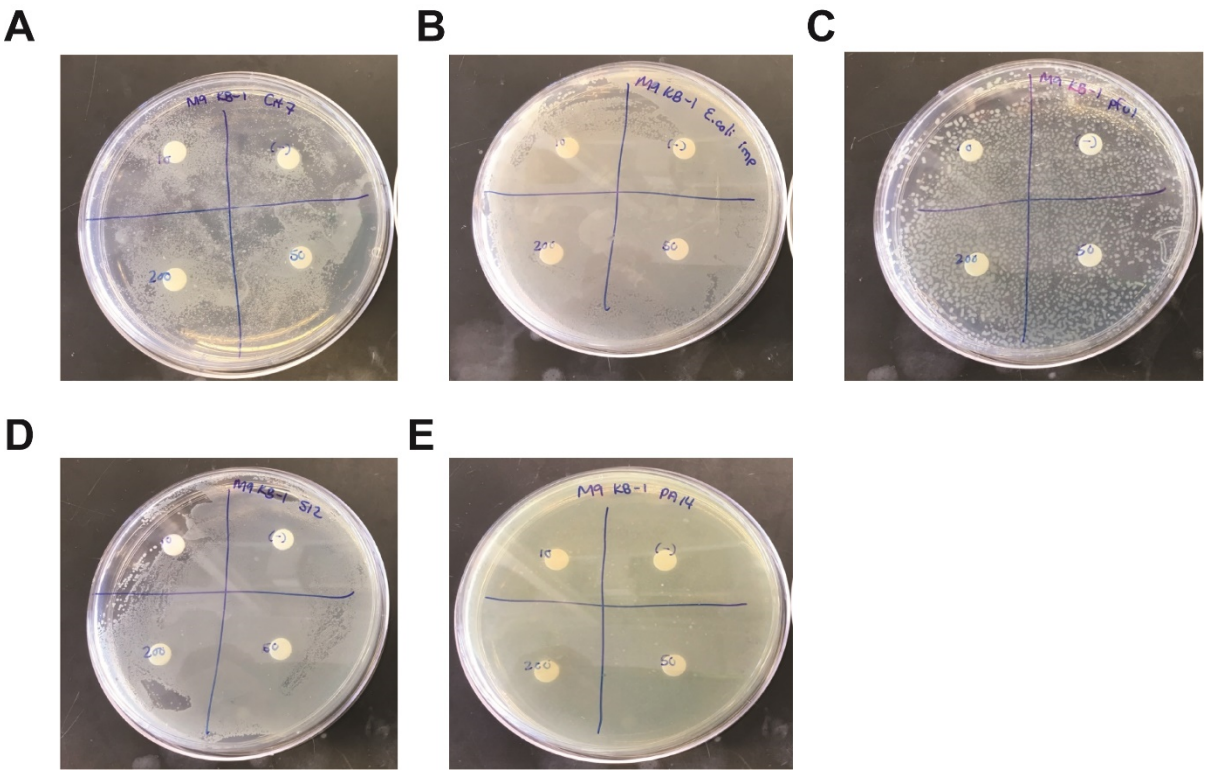


Figure S3.1. Agar diffusion assays for KB-1. Varying concentrations of KB-1 were plated on filter paper atop M9 agar media plated with (*A) P. syringae* Cit7*, (B) E. coli* MG1655*, (C) P. fluorescens* Pf0-1*, (D) P. putida* S12, and *(E) P. aeruginosa* PA14. No zones of inhibition were observed for any strains tested.

REFERENCES

(1) Fawaz, M. V.; Topper, M. E.; Firestine, S. M., The ATP-grasp enzymes. *Bioorg Chem* **2011,** *39*, 185-191.

(2) Wang, W.; Kappock, T. J.; Stubbe, J.; Ealick, S. E., X-ray crystal structure of glycinamide ribonucleotide synthetase from Escherichia coli. *Biochemistry* **1998,** *37* (45), 15647-15662.

(3) Goswami, A.; Van Lanen, S. G., Enzymatic strategies and biocatalysts for amide bond formation: tricks of the trade outside of the ribosome. *Mol Biosyst* **2015,** *11*, 338-353.

(4) Iyer, L. M.; Abhiman, S.; Burroughs, A. M.; Aravind, L., Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol Biosyst* **2009,** *5*, 1636-1660.

(5) Chou, C.-Y.; Linda, P.; Tong, L., Crystal structure of biotin carboxylase in complex with substrates and implications for its catalytic mechanism. *Journal of Biological Chemistry* **2009**.

(6) Li, H.; Fast, W.; Benkovic, S. J., Structural and functional modularity of proteins in the de novo purine biosynthetic pathway. *Protein Science* **2009,** *18* (5), 881-892.

(7) Galperin, M. Y.; Koonin, E. V., A diverse superfamily of enzymes with ATP-dependent carboxylate—amine/thiol ligase activity. *Protein Science* **1997,** *6* (12), 2639-2643.

(8) Zhao, G.; Jin, Z.; Wang, Y.; Allewell, N. M.; Tuchman, M.; Shi, D., Structure and function of Escherichia coli RimK, an ATP-grasp fold, l-glutamyl ligase enzyme. *Proteins: Structure, Function, and Bioinformatics* **2013,** *81* (10), 1847-1854.

(9) Lee, D.; Redfern, O.; Orengo, C., Predicting protein function from sequence and structure. *Nature Reviews Molecular Cell Biology* **2007,** *8* (12), 995.

(10) Arai, T.; Arimura, Y.; Ishikura, S.; Kino, K., L-Amino acid ligase from Pseudomonas syringae producing tabtoxin can be used for the enzymatic synthesis of various functional peptides. *Applied and environmental microbiology* **2013**, AEM. 01003-13.

(11) Arai, T.; Kino, K., A novel L-amino acid ligase is encoded by a gene in the phaseolotoxin biosynthetic gene cluster from *Pseudomonas syringae* pv. phaseolicola 1448A. *Biosci Biotechnol Biochem* **2008,** *72*, 3048-3050.

(12) O'Neill, E. M.; Mucyn, T. S.; Patteson, J. B.; Finkel, O. M.; Chung, E.-H.; Baccile, J. A.; Massolo, E.; Schroeder, F. C.; Dangl, J. L.; Li, B., Phevamine A, a small molecule that suppresses plant immune responses. *Proceedings of the National Academy of Sciences* **2018,** *115* (41), E9514-E9522.

(13) Mucyn, T. S.; Yourstone, S.; Lind, A. L.; Biswas, S.; Nishimura, M. T.; Baltrus, D. A.; Cumbie, J. S.; Chang, J. H.; Jones, C. D.; Dangl, J. L., Variable suites of non-effector genes

are co-regulated in the type III secretion virulence regulon across the *Pseudomonas syringae* phylogeny. *PLoS Pathog* **2014,** *10*, e1003807.

(14) Carrión, V. J.; Arrebola, E.; Cazorla, F. M.; Murillo, J.; de Vicente, A., The mbo operon is specific and essential for biosynthesis of mangotoxin in *Pseudomonas syringae*. *PLoS One* **2012,** *7*, e36709.

(15) Arrebola, E.; Cazorla, F. M.; Durán, V. E.; Rivera, E.; Olea, F.; Codina, J. C.; Pérez-Garcıa, A.; de Vicente, A., Mangotoxin: a novel antimetabolite toxin produced by Pseudomonas syringae inhibiting ornithine/arginine biosynthesis. *Physiological and Molecular Plant Pathology* **2003,** *63* (3), 117-127.

(16) Hollenhorst, M. A.; Clardy, J.; Walsh, C. T., The ATP-dependent amide ligases DdaG and DdaF assemble the fumaramoyl-dipeptide scaffold of the dapdiamide antibiotics. *Biochemistry* **2009,** *48*, 10467-10472.

(17) Uda, N.; Matoba, Y.; Kumagai, T.; Oda, K.; Noda, M.; Sugiyama, M., Establishment of an in vitro D-cycloserine-synthesizing system by using O-ureido-L-serine synthase and D-cycloserine synthetase found in the biosynthetic pathway. *Antimicrobial agents and chemotherapy* **2013**, AAC. 02291-12.

(18) Gao, Q.; Garcia-Pichel, F., An ATP-grasp ligase involved in the last biosynthetic step of the imino-mycosporine, shinorine, in *Nostoc punctiforme* ATCC29133. *Journal of bacteriology* **2011**, JB. 05730-11.

(19) Ju, K.-S.; Gao, J.; Doroghazi, J. R.; Wang, K.-K. A.; Thibodeaux, C. J.; Li, S.; Metzger, E.; Fudala, J.; Su, J.; Zhang, J. K., Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proceedings of the National Academy of Sciences* **2015,** *112* (39), 12175-12180.

(20) Lauer, B.; Russwurm, R.; Schwarz, W.; Kalmanczhelyi, A.; Bruntner, C.; Rosemeier, A.; Bormann, C., Molecular characterization of co-transcribed genes from Streptomyces tendae Tü901 involved in the biosynthesis of the peptidyl moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. *Molecular and General Genetics MGG* **2001,** *264* (5), 662-673.

(21) Schneider, T.; Gries, K.; Josten, M.; Wiedemann, I.; Pelzer, S.; Labischinski, H.; Sahl, H.-G., The lipopeptide antibiotic Friulimicin B inhibits cell wall biosynthesis through complex formation with bactoprenol phosphate. *Antimicrobial agents and chemotherapy* **2009,** *53* (4), 1610-1618.

(22) Borisova, S. A.; Circello, B. T.; Zhang, J. K.; van der Donk, W. A.; Metcalf, W. W., Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by Bacillus subtilis ATCC6633. *Chemistry & biology* **2010,** *17* (1), 28-37.

(23) Metcalf, W. W.; van der Donk, W. A.; Zhang, J.; Circello, B. T.; Borisova, S. A., Compositions and methods for the synthesis of APPA-containing peptides. Google Patents: 2013.

(24) Gao, J.; Ju, K. S.; Yu, X.; Velásquez, J. E.; Mukherjee, S.; Lee, J.; Zhao, C.; Evans, B. S.; Doroghazi, J. R.; Metcalf, W. W., Use of a phosphonate methyltransferase in the identification of the fosfazinomycin biosynthetic gene cluster. *Angewandte Chemie* **2014,** *126* (5), 1358-1361.

(25) Schiessl, K.; Roller, A.; Hammerschmidt, F., Determination of absolute configuration of the phosphonic acid moiety of fosfazinomycins. *Organic & biomolecular chemistry* **2013,** *11* (42), 7420-7426.

(26) Gunji, S.; Arima, K.; Beppu, T., Screening of antifungal antibiotics according to activities inducing morphological abnormalities. *Agricultural and Biological Chemistry* **1983,** *47* (9), 2061-2069.

(27) Li, K.; Condurso, H. L.; Li, G.; Ding, Y.; Bruner, S. D., Structural basis for precursor protein–directed ribosomal peptide macrocyclization. *Nature chemical biology* **2016,** *12* (11), 973.

(28) Noike, M.; Matsui, T.; Ooya, K.; Sasaki, I.; Ohtaki, S.; Hamano, Y.; Maruyama, C.; Ishikawa, J.; Satoh, Y.; Ito, H., A peptide ligase and the ribosome cooperate to synthesize the peptide pheganomycin. *Nature chemical biology* **2015,** *11* (1), 71.

(29) Tsuda, T.; Suzuki, T.; Kojima, S., Crystallization and preliminary X-ray diffraction analysis of Bacillus subtilis YwfE, an l-amino-acid ligase. *Acta crystallographica section F: Structural biology and crystallization communications* **2012,** *68* (2), 203-206.

(30) Suzuki, M.; Takahashi, Y.; Noguchi, A.; Arai, T.; Yagasaki, M.; Kino, K.; Saito, J. i., The structure of L-amino-acid ligase from Bacillus licheniformis. *Acta Crystallographica Section D: Biological Crystallography* **2012,** *68* (11), 1535-1540.

(31) Tabata, K.; Ikeda, H.; Hashimoto, S.-i., ywfE in *Bacillus subtilis* codes for a novel enzyme, l-amino acid ligase. *Journal of bacteriology* **2005,** *187* (15), 5195-5202.

(32) Shomura, Y.; Hinokuchi, E.; Ikeda, H.; Senoo, A.; Takahashi, Y.; Saito, J. i.; Komori, H.; Shibata, N.; Yonetani, Y.; Higuchi, Y., Structural and enzymatic characterization of BacD, an L-amino acid dipeptide ligase from Bacillus subtilis. *Protein science* **2012,** *21* (5), 707-716.

(33) Steinborn, G.; Hajirezaei, M.-R.; Hofemeister, J., bac genes for recombinant bacilysin and anticapsin production in Bacillus host strains. *Archives of microbiology* **2005,** *183* (2), 71-79.

(34) Kenig, M.; Vandamme, E.; Abraham, E., The mode of action of bacilysin and anticapsin and biochemical properties of bacilysin-resistant mutants. *Microbiology* **1976,** *94* (1), 46-54.

(35) Kino, K.; Arai, T.; Tateiwa, D., A novel L-amino acid ligase from Bacillus subtilis NBRC3134 catalyzed oligopeptide synthesis. *Bioscience, biotechnology, and biochemistry* **2010,** *74* (1), 129-134.

(36) Kino, K.; Kotanaka, Y.; Arai, T.; Yagasaki, M., A novel L-amino acid ligase from Bacillus subtilis NBRC3134, a microorganism producing peptide-antibiotic rhizocticin. *Bioscience, biotechnology, and biochemistry* **2009,** *73* (4), 901-907.

(37) Kagawa, W.; Arai, T.; Ishikura, S.; Kino, K.; Kurumizaka, H., Structure of RizA, an l-amino-acid ligase from Bacillus subtilis. *Acta Crystallographica Section F: Structural Biology Communications* **2015,** *71* (9), 1125-1130.

(38) Blasiak, L. C.; Clardy, J., Discovery of 3-formyl-tyrosine metabolites from *Pseudoalteromonas tunicata* through heterologous expression. *J Am Chem Soc* **2009,** *132*, 926-927.

(39) Bender, C. L.; Alarcón-Chaidez, F.; Gross, D. C., *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol Mol Biol R* **1999,** *63*, 266-292.

(40) Katagiri, F.; Thilmony, R.; He, S. Y., The *Arabidopsis thaliana-Pseudomonas syringae* interaction. *The Arabidopsis Book* **2002**, e0039.

(41) Geng, X.; Jin, L.; Shimada, M.; Kim, M. G.; Mackey, D., The phytotoxin coronatine is a multifunctional component of the virulence armament of *Pseudomonas syringae*. *Planta* **2014,** *240*, 1149-1165.

(42) Ferguson, A.; Johnston, J., Phaseolotoxin: chlorosis, ornithine accumulation and inhibition of ornithine carbamoyltransferase in different plants. *Physiol Plant Pathol* **1980,** *16*, 269-275.

(43) Bachmann, A.; Matile, P.; Slusarenko, A., Inhibition of ornithine decarboxylase activity by phaseolotoxin: implications for symptom production in halo blight of French bean. *Physiol Mol Plant P* **1998,** *53*, 287-299.

(44) Gulick, A. M., Nonribosomal peptide synthetase biosynthetic clusters of ESKAPE pathogens. *Nat Prod Rep* **2017,** *34*, 981-1009.

(45) Gross, H.; Loper, J. E., Genomics of secondary metabolite production by *Pseudomonas spp*. *Nat Prod Rep* **2009,** *26*, 1408-1446.

(46) Gimenez-Ibanez, S.; Chini, A.; Solano, R., How microbes twist jasmonate signaling around their little fingers. *Plants* **2016,** *5*, 9.

(47) Toruño, T. Y.; Stergiopoulos, I.; Coaker, G., Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. *Annu Rev Phytopathol* **2016,** *54*, 419-441.

(48) Yu, X.; Feng, B.; He, P.; Shan, L., From chaos to harmony: Responses and signaling upon microbial pattern recognition. *Annu Rev Phytopathol* **2017,** *55*, 109-137.

(49) Fouts, D. E.; Abramovitch, R. B.; Alfano, J. R.; Baldo, A. M.; Buell, C. R.; Cartinhour, S.; Chatterjee, A. K.; D'Ascenzo, M.; Gwinn, M. L.; Lazarowitz, S. G., Genomewide identification of *Pseudomonas syringae* pv. tomato DC3000 promoters controlled by the HrpL alternative sigma factor. *Proc Natl Acad Sci USA* **2002,** *99*, 2275-2280.

(50) Weingart, H.; Stubner, S.; Schenk, A.; Ullrich, M. S., Impact of temperature on in planta expression of genes involved in synthesis of the *Pseudomonas syringae* phytotoxin coronatine. *Mol Plant Microbe Interact* **2004,** *17*, 1095-1102.

(51) Lan, L.; Deng, X.; Zhou, J.; Tang, X., Genome-wide gene expression analysis of *Pseudomonas syringae* pv. tomato DC3000 reveals overlapping and distinct pathways regulated by *hrp*L and *hrp*RS. *Mol Plant Microbe Interact* **2006,** *19*, 976-987.

(52) Oh, C.-S.; Beer, S. V., Molecular genetics of *Erwinia amylovora* involved in the development of fire blight. *FEMS Microbiol Lett* **2005,** *253*, 185-192.

(53) Oh, C. S.; Kim, J. F.; Beer, S. V., The Hrp pathogenicity island of *Erwinia amylovora* and identification of three novel genes required for systemic infection. *Mol Plant Pathol* **2005,** *6*, 125-138.

(54) Kvitko, B. H.; Park, D. H.; Velásquez, A. C.; Wei, C.-F.; Russell, A. B.; Martin, G. B.; Schneider, D. J.; Collmer, A., Deletions in the repertoire of *Pseudomonas syringae* pv. tomato DC3000 type III secretion effector genes reveal functional overlap among effectors. *PLoS Pathogens* **2009,** *5*, e1000388.

(55) Gómez-Gómez, L.; Felix, G.; Boller, T., A single locus determines sensitivity to bacterial flagellin in *Arabidopsis thaliana*. *Plant J* **1999,** *18*, 277-284.

(56) Bowles, D. J., Defense-related proteins in higher plants. *Annu Rev Biochem* **1990,** *59*, 873-907.

(57) Hauck, P.; Thilmony, R.; He, S. Y., A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible *Arabidopsis* plants. *Proc Natl Acad Sci USA* **2003,** *100*, 8577-8582.

(58) DebRoy, S.; Thilmony, R.; Kwack, Y.-B.; Nomura, K.; He, S. Y., A family of conserved bacterial effectors inhibits salicylic acid-mediated basal immunity and promotes disease necrosis in plants. *Proc Natl Acad Sci USA* **2004,** *101*, 9927-9932.

(59) De Torres, M.; Mansfield, J. W.; Grabov, N.; Brown, I. R.; Ammouneh, H.; Tsiamis, G.; Forsyth, A.; Robatzek, S.; Grant, M.; Boch, J., *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in *Arabidopsis*. *Plant J* **2006,** *47*, 368-382.

(60) Nomura, K.; DebRoy, S.; Lee, Y. H.; Pumplin, N.; Jones, J.; He, S. Y., A bacterial virulence protein suppresses host innate immunity to cause plant disease. *Science* **2006,** *313*, 220-223.

(61) Underwood, W.; Zhang, S.; He, S. Y., The *Pseudomonas syringae* type III effector tyrosine phosphatase HopAO1 suppresses innate immunity in *Arabidopsis thaliana*. *Plant J* **2007,** *52*, 658-672.

(62) Cunnac, S.; Chakravarthy, S.; Kvitko, B. H.; Russell, A. B.; Martin, G. B.; Collmer, A., Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. *Proc Natl Acad Sci USA* **2011,** *108*, 2975-2980.

(63) Silby, M. W.; Cerdeño-Tárraga, A. M.; Vernikos, G. S.; Giddens, S. R.; Jackson, R. W.; Preston, G. M.; Zhang, X.-X.; Moon, C. D.; Gehrig, S. M.; Godfrey, S. A., Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biol* **2009,** *10*, R51.

(64) Thomas, W. J.; Thireault, C. A.; Kimbrel, J. A.; Chang, J. H., Recombineering and stable integration of the *Pseudomonas syringae* pv. syringae 61 hrp/hrc cluster into the genome of the soil bacterium *Pseudomonas fluorescens Pf*0-1. *Plant J* **2009,** *60*, 919-928.

(65) Oh, H.-S.; Park, D. H.; Collmer, A., Components of the *Pseudomonas syringae* type III secretion system can suppress and may elicit plant innate immunity. *Mol Plant Microbe In* **2010,** *23*, 727-739.

(66) He, S. Y.; Huang, H.-C.; Collmer, A., *Pseudomonas syringae* pv. syringae harpinPss: a protein that is secreted via the Hrp pathway and elicits the hypersensitive response in plants. *Cell* **1993,** *73*, 1255-1266.

(67) Charkowski, A. O.; Alfano, J. R.; Preston, G.; Yuan, J.; He, S. Y.; Collmer, A., The *Pseudomonas syringae* pv. tomato HrpW protein has domains similar to harpins and pectate lyases and can elicit the plant hypersensitive response and bind to pectate. *J Bacteriol* **1998,** *180*, 5211-5217.

(68) Michael, A. J.; Furze, J. M.; Rhodes, M. J.; Burtin, D., Molecular cloning and functional identification of a plant ornithine decarboxylase cDNA. *Biochem J* **1996,** *314*, 241-248.

(69) Hamana, K.; Sakamoto, A.; Tachiyanagi, S.; Terauchi, E., Polyamine profiles of some members of the gamma subclass of the class Proteobacteria: polyamine analysis of twelve recently described genera. *Microbiol Cult Collect* **2003,** *19*, 3-11.

(70) Medema, M. H.; Takano, E.; Breitling, R., Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* **2013,** *30*, 1218-1223.

(71) Baltrus, D. A.; Nishimura, M. T.; Romanchuk, A.; Chang, J. H.; Mukhtar, M. S.; Cherkis, K.; Roach, J.; Grant, S. R.; Jones, C. D.; Dangl, J. L., Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog* **2011,** *7*, e1002132.

(72) McCann, H. C.; Rikkerink, E. H.; Bertels, F.; Fiers, M.; Lu, A.; Rees-George, J.; Andersen, M. T.; Gleave, A. P.; Haubold, B.; Wohlers, M. W., Genomic analysis of the

kiwifruit pathogen *Pseudomonas syringae* pv. actinidiae provides insight into the origins of an emergent plant disease. *PLoS Pathog* **2013,** *9*, e1003503.

(73) Yang, L.; Teixeira, P. J. P. L.; Biswas, S.; Finkel, O. M.; He, Y.; Salas-Gonzalez, I.; English, M. E.; Epple, P.; Mieczkowski, P.; Dangl, J. L., *Pseudomonas syringae* Type III effector HopBB1 promotes host transcriptional repressor degradation to regulate phytohormone responses and virulence. *Cell Host Microbe* **2017,** *21*, 156-168.

(74) Mitchell, R. E., Isolation and structure of a chlorosis-inducing toxin of *Pseudomonas phaseolicola*. *Phytochemistry* **1976,** *15*, 1941-1947.

(75) Hussain, S. S.; Ali, M.; Ahmad, M.; Siddique, K. H., Polyamines: natural and engineered abiotic and biotic stress tolerance in plants. *Biotechnol Adv* **2011,** *29*, 300-311.

(76) Jiménez Bremont, J. F.; Marina, M.; Guerrero-González, M. d. l. L.; Rossi, F. R.; Sánchez-Rangel, D.; Rodríguez-Kessler, M.; Ruiz, O. A.; Gárriz, A., Physiological and molecular implications of plant polyamine metabolism during biotic interactions. *Front Plant Sci* **2014,** *5*, 95.

(77) Carrión, V. J.; Gutiérrez-Barranquero, J. A.; Arrebola, E.; Bardaji, L.; Codina, J. C.; de Vicente, A.; Cazorla, F. M.; Murillo, J., The mangotoxin biosynthetic operon (mbo) is specifically distributed within *Pseudomonas syringae* genomospecies 1 and was acquired only once during evolution. *Appl Environ Microb* **2013,** *79*, 756-767.

(78) Zipfel, C.; Robatzek, S.; Navarro, L.; Oakeley, E. J.; Jones, J. D.; Felix, G.; Boller, T., Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* **2004,** *428*, 764-767.

(79) Chinchilla, D.; Bauer, Z.; Regenass, M.; Boller, T.; Felix, G., The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *Plant Cell* **2006,** *18*, 465-476.

(80) Kunze, G.; Zipfel, C.; Robatzek, S.; Niehaus, K.; Boller, T.; Felix, G., The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *Plant Cell* **2004,** *16*, 3496-3507.

(81) Yoda, H.; Fujimura, K.; Takahashi, H.; Munemura, I.; Uchimiya, H.; Sano, H., Polyamines as a common source of hydrogen peroxide in host-and nonhost hypersensitive response during pathogen infection. *Plant Mol Biol* **2009,** *70*, 103-112.

(82) Ward, J. L.; Forcat, S.; Beckmann, M.; Bennett, M.; Miller, S. J.; Baker, J. M.; Hawkins, N. D.; Vermeer, C. P.; Lu, C.; Lin, W., The metabolic transition during disease following infection of *Arabidopsis thaliana* by *Pseudomonas syringae* pv. tomato. *Plant J* **2010,** *63*, 443-457.

(83) Lou, Y.-R.; Bor, M.; Yan, J.; Preuss, A. S.; Jander, G., *Arabidopsis* NATA1 acetylates putrescine and decreases defense-related hydrogen peroxide accumulation. *Plant Physiol* **2016**, pp. 00446.2016.

(84) Gonzalez, M. E.; Marco, F.; Minguet, E. G.; Sorli, P. C.; Blázquez, M. A.; Carbonell, J.; Ruiz, O. A.; Pieckenstain, F. L., Perturbation of spermine synthase gene expression and transcript profiling provide new insights on the role of the tetraamine spermine in *Arabidopsis thaliana* defense against *Pseudomonas viridiflava*. *Plant Physiol* **2011**, pp. 110.171413.

(85) Knight, H.; Trewavas, A. J.; Knight, M. R., Cold calcium signaling in *Arabidopsis* involves two cellular pools and a change in calcium signature after acclimation. *Plant Cell* **1996,** *8*, 489-503.

(86) Nelson, J. K.; Frølund, S. U.; Tikhonov, D. B.; Kristensen, A. S.; Strømgaard, K., Synthesis and biological activity of argiotoxin 636 and analogues: selective antagonists for ionotropic glutamate receptors. *Angew Chem Int Ed* **2009,** *48*, 3087-3091.

(87) Ogasawara, Y.; Dairi, T., Biosynthesis of Oligopeptides using ATP-grasp Enzymes. *Chem-Eur J* **2017,** *23*, 10714-10724.

(88) Ogasawara, Y.; Kawata, J.; Noike, M.; Satoh, Y.; Furihata, K.; Dairi, T., Exploring peptide ligase orthologs in *Actinobacteria*: Discovery of pseudopeptide natural products, ketomemicins. *ACS Chem Biol* **2016,** *11*, 1686-1692.

(89) Sambrook, J.; Russell, D., *Molecular Cloning: A Laboratory Manual*. 3rd ed.; Cold Spring Harbor: New York, 2001; p 6.4-6.12.

(90) Tornero, P.; Dangl, J. L., A high-throughput method for quantifying growth of phytopathogenic bacteria in *Arabidopsis thaliana*. *Plant J* **2001,** *28*, 475-481.

(91) Kim, M. G.; Da Cunha, L.; McFall, A. J.; Belkhadir, Y.; DebRoy, S.; Dangl, J. L.; Mackey, D., Two *Pseudomonas syringae* type III effectors inhibit RIN4-regulated basal defense in *Arabidopsis*. *Cell* **2005,** *121*, 749-759.

(92) Letunic, I.; Bork, P., Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **2016,** *44*, W242-W245.

(93) Obranić, S.; Babić, F.; Maravić-Vlahoviček, G., Improvement of pBBR1MCS plasmids, a very useful series of broad-host-range cloning vectors. *Plasmid* **2013,** *70*, 263-267.

(94) Choi, K.-H.; Schweizer, H. P., mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nat Protoc* **2006,** *1*, 153.

(95) Marco, M. L.; Legac, J.; Lindow, S. E., *Pseudomonas syringae* genes induced during colonization of leaf surfaces. *Environ Microbiol* **2005,** *7*, 1379-1391.

(96) Reyrat, J.-M.; Pelicic, V.; Gicquel, B.; Rappuoli, R., Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infect Immun* **1998,** *66*, 4011-4017.

(97) Kaur, N.; Delcros, J.-G.; Archer, J.; Weagraff, N. Z.; Martin, B.; Phanstiel Iv, O., Designing the polyamine pharmacophore: Influence of N-substituents on the transport behavior of polyamine conjugates. *J Med Chem* **2008,** *51*, 2551-2560.

(98) Wang, J.; Gao, R.; Li, Q.; Xie, S.; Zhao, J.; Wang, C., Synthesis, cytotoxicity, and cell death profile of polyaminoanthraquinones as antitumor agents. *Chem Biol Drug Des* **2012,** *80*, 909-917.

(99) Wang, J.; Gan, Y.; Li, S.; Luo, T.; Zhang, Y.; Zhao, J., Potent P-glycoprotein inhibition of emodin derivative: synthesis and biological evaluation. *Med Chem Res* **2014,** *23*, 2106-2112.

(100) Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L., Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics* **2015,** *1854* (8), 1019-1037.

(101) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **2003,** *13* (11), 2498-2504.

(102) Smoot, M. E.; Ono, K.; Ruscheinski, J.; Wang, P.-L.; Ideker, T., Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **2010,** *27* (3), 431-432.

(103) Graniti, A., Phytotoxins and their involvement in plant diseases. Introduction. *Experientia* **1991,** *47* (8), 751-755.

(104) Lamichhane, J. R.; Varvaro, L.; Parisi, L.; Audergon, J.-M.; Morris, C. E., Disease and frost damage of woody plants caused by *Pseudomonas syringae*: seeing the forest for the trees. In *Advances in agronomy*, Elsevier: 2014; Vol. 126, pp 235-295.

(105) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H., Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evolutionary Biology* **2007,** *7* (1), 78.

(106) Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L.; Marshall, C. G.; Miller, D. A.; Patel, H. M., Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Current opinion in chemical biology* **2001,** *5* (5), 525-534.

(107) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O., NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. *Nucleic acids research* **2011,** *39* (suppl_2), W362-W367.

(108) Skinnider, M. A.; Dejong, C. A.; Rees, P. N.; Johnston, C. W.; Li, H.; Webster, A. L.; Wyatt, M. A.; Magarvey, N. A., Genomes to natural products prediction informatics for secondary metabolomes (PRISM). *Nucleic acids research* **2015,** *43* (20), 9645-9662.

(109) Hubbard, B. K.; Thomas, M. G.; Walsh, C. T., Biosynthesis of Lp-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. *Chemistry & biology* **2000,** *7* (12), 931-942.

(110) Baltz, R. H., Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. *Journal of industrial microbiology & biotechnology* **2011,** *38* (11), 1747.

(111) Felnagle, E. A.; Barkei, J. J.; Park, H.; Podevels, A. M.; McMahon, M. D.; Drott, D. W.; Thomas, M. G., MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. *Biochemistry* **2010,** *49* (41), 8815-8817.

(112) Li, C.; Kelly, W. L., Recent advances in thiopeptide antibiotic biosynthesis. *Natural product reports* **2010,** *27* (2), 153-164.

(113) Maier, S.; Heitzler, T.; Asmus, K.; Brötz, E.; Hardter, U.; Hesselbach, K.; Paululat, T.; Bechthold, A., Functional Characterization of Different ORFs Including Luciferase-Like Monooxygenase Genes from the Mensacarcin Gene Cluster. *ChemBioChem* **2015,** *16* (8), 1175-1182.

(114) Thutewohl, M.; Waldmann, H., Solid-phase synthesis of a pepticinnamin E library. *Bioorganic & medicinal chemistry* **2003,** *11* (12), 2591-2615.

(115) Hinterding, K.; Hagenbuch, P.; Rétey, J.; Waldmann, H., Synthesis and in vitro evaluation of the Ras farnesyltransferase inhibitor pepticinnamin E. *Angewandte Chemie International Edition* **1998,** *37* (9), 1236-1239.

(116) Singh, D.; Tripathi, A.; Kumar, G., An Overview of Computational Approaches in Structure Based Drug Design. *Nepal Journal of Biotechnology* **2012,** *2* (1), 53-61.

(117) Shiomi, K.; Yang, H.; Inokoshi, J.; Van Der Pyl, D.; Nakagawa, A.; Takeshima, H.; Omura, S., Pepticinnamins, new farnesyl-protein transferase inhibitors produced by an actinomycete. *The Journal of antibiotics* **1993,** *46* (2), 229-234.

(118) Omura, S.; Van Der Pyl, D.; Inokoshi, J.; Takahashi, Y.; Takeshima, H., Pepticinnamins, new farnesyl-protein transferase inhibitors produced by an actinomycete. *The Journal of antibiotics* **1993,** *46* (2), 222-228.

(119) Sun, D., Highly stereoselective and efficient synthesis of the dopa analogue in pepticinnamin E via enantioselective hydrogenation of dehydroamino acids. *Turkish Journal of Chemistry* **2010,** *34* (2), 181-186.

(120) Hamdan, F. F.; Ribeiro, P., Cloning and characterization of a novel form of tyrosine hydroxylase from the human parasite, Schistosoma mansoni. *Journal of neurochemistry* **1998,** *71* (4), 1369-1380.

(121) Pomerantz, S. H., The tyrosine hydroxylase activity of mammalian tyrosinase. *Journal of Biological Chemistry* **1966,** *241* (1), 161-168.

(122) Zhang, W.; Ames, B. D.; Walsh, C. T., Identification of phenylalanine 3-hydroxylase for meta-tyrosine biosynthesis. *Biochemistry* **2011,** *50* (24), 5401-5403.

(123) Nagatsu, T.; Oka, K.; Kato, T., Highly sensitive assay for tyrosine hydroxylase activity by high-performance liquid chromatography. *Journal of Chromatography B: Biomedical Sciences and Applications* **1979,** *163* (3), 247-252.

(124) Ma, S.; Morris, V.; Cuppels, D., Characterization of a DNA region required for production of the phytotoxin coronatine by *Pseudomonas syringae* pv. tomato. *Mol Plant Microbe Interact* **1991,** *4*, 69-74.

(125) Zeng, W.; Brutus, A.; Kremer, J. M.; Withers, J. C.; Gao, X.; Jones, A. D.; He, S. Y., A genetic screen reveals *Arabidopsis* stomatal and/or apoplastic defenses against *Pseudomonas syringae* pv. tomato DC3000. *PLoS Pathog* **2011,** *7*, e1002291.

(126) Carrión, V. J.; Arrebola, E.; Cazorla, F. M.; Murillo, J.; de Vicente, A., The mbo operon is specific and essential for biosynthesis of mangotoxin in *Pseudomonas syringae*. *PLoS One* **2012,** *7* (5), e36709.

(127) Klingeman, D. M.; Utturkar, S.; Lu, T.-Y. S.; Schadt, C. W.; Pelletier, D. A.; Brown, S. D., Draft genome sequences of four *Streptomyces* isolates from the *Populus trichocarpa* root endosphere and rhizosphere. *Genome announcements* **2015,** *3* (6), e01344-15.

(128) Yamasaki, H.; Cohen, M. F., NO signal at the crossroads: polyamine-induced nitric oxide synthesis in plants? *Trends Plant Sci* **2006,** *11*, 522-524.

(129) Kim, N. H.; Kim, B. S.; Hwang, B. K., Pepper arginine decarboxylase is required for polyamine and γ-aminobutyric acid signaling in cell death and defense response. *Plant Physiol* **2013,** *162*, 2067-2083.